# Chemical Approaches for Advanced Optical Imaging

Zhixing Chen

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT

# Chemical Approaches for Advanced Optical Imaging

Zhixing Chen

Advances in optical microscopy have been constantly expanding our knowledge of biological systems. The achievements therein are a result of close collaborations between physicists/engineers who build the imaging instruments and chemists/biochemists who design the corresponding probe molecules. In this work I present a number of chemical approaches for the development of advanced optical imaging methods. Chapter 1 provides an overview of the recent advances of novel imaging approaches taking advantage of chemical tag technologies. Chapter 2 describes the second-generation covalent trimethoprim-tag as a viable tool for live cell protein-specific labeling and imaging. In Chapter 3 we present a fluorescence lifetime imaging approach to map protein-specific micro-environment in live cells using TMP-Cy3 as a chemical probe. In Chapter 4, we present a method harnessing photo-activatable fluorophores to extend the fundamental depth limit in multi-photon microscopy. Chapter 5 describes the development of isotopically edited alkyne palette for multi-color live cell vibrational imaging of cellular small molecules. These studies exemplify the impact of modern chemical approaches in the development of advanced optical microscopies.

## Table of Contents

## List of Figures

## List of Tables

## List of Schemes

**Acknowledgments**

I would first like to give my thanks to Professors Virginia Cornish and Wei Min, for their mentorships to my graduate study. I feel extremely fortunate to have the two enlightening advisors. It was Virginia's passion on interdisciplinary science which expanded my research interest from organic chemistry to chemical biology. And it was Wei's deep commitment on microscopy which inspired me to work on projects related to both chemistry and optical imaging. The scientifically complementary yet personally collaborative nature of the two labs have rendered me unique research opportunities. And it is Virginia and Wei who made it possible. Beside the sciences, I am heavily influenced by the "people-oriented philosophy" from Virginia, I admire Wei's unique sense to theories, and from both Virginia and Wei I got a brave attitude towards emerging scientific adventures. These aspects of their characters have become an invaluable heritage to me.

I would like to thank the faculty of Columbia University, who have established a wonderful community for scientific research. I thank Professors Scott Snyder and Tristan Lambert for recruiting me to Columbia. I thank Professors Michael Sheetz, Colin Nuckolls and Richard Friesner for their shares to the collaborative research projects I have been carrying on. I thank Professors Jonathan Owen, Ruben Gonzalez, Elizabeth Hillman, James Hone, Martin Chalfie and Rafael Yuste for the discussions on my research. I give special thanks to Professors Ronald Breslow, Dalibor Sames and Howard Hang for critically serving as my thesis examiners.

I would like to thank all the collaborators on my projects at Columbia. Specifically, Chaoran Jing for her initial help when I first joined Columbia; Andrew Anzalone for his enjoyable approach to oxazine chemistry; Miguel Jimenez on his discussions and proofreadings of my projects; Lu Wei for her entrepreneurial care to the projects in Min lab. I would also like to thank Tracy Wang, Evangelos Gatzogiannis, Daniel Paley, Andrew Weisman, Yihui Shen, Fanghao Hu, Fang Xu, Annie Lin, Xinxin Zhu, Richard Wombacher, Ya-Ting Kao, Luyuan Zhang, Rachel Fleisher, Bertrand Adanve, Nili Ostrov, Mia Shandell and Andy Ng. It is hard to exhaust all, but it would never be possible for me to make the projects happen without the collaborators' contributions, scientifically and personally.

I would like to thank all my mentors and friends in science for the supports along my journey. It was Professor Judong Cao, a dedicated chemistry educator who opened the door of chemistry to me. It was Professor Zhen Yang at Peking University who brought me to modern research by a fascinating molecule - a target of total synthesis. It is Boyuan Wang who have been fighting together with me as chemistry students. I would like to thank Professors Lei Liu, David Zhigang Wang and Jinying Yuan for the encouraging discussions which finally led me into graduate school. I would also thank Xiaochun Li and Weixian Xi as college roommates who have been pursuing science together with me.

Finally I would like to thank my parents, who have always been supportive to let me pursue my own dreams. I thank Li An, my girlfriend, who showed up late enough in my life so that I can finish my thesis while early enough so we can share the moment.

# Chapter 1

# Chemical tags: inspiration for advanced imaging techniques

The content of this chapter has been published in:

Chen, Z.; Cornish, V. W.; and Min, W. “Chemical tags: inspiration for advanced imaging techniques”. *Curr. Opin. Chem. Biol.* **2013**, 17, 637.

## 1.1 Chapter outlook

This chapter summarizes recent applications of chemical tags in conjunction with advanced bio-imaging techniques including single-molecule fluorescence, spatiotemporally resolved ensemble microscopy techniques, and imaging modalities beyond fluorescence. We aim to illustrate the unique advantages of chemical tags in facilitating contemporary microscopy to address biological problems that are difficult or near impossible to approach otherwise. We hope this chapter will inspire more innovative applications enabled by the mingling of these two growing fields.

I have been the main contributor of this chapter. I wrote the chapter with contributions from Wei Min and Virginia Cornish.

## 1.2 Introduction

Advances in microscopy have tremendously expanded our knowledge of biological processes at the microscopic level. The achievements therein are the result of close collaborations between physicists/engineers who build the imaging instruments and chemists/biochemists who design the corresponding probe molecules. One classic example representing this trend is the use of GFP to visualize specific proteins within living organisms by fluorescence microscopy[1]. Recent developments in more advanced imaging schemes (e.g. single-molecule fluorescence imaging, fluorescence lifetime imaging, triplet-state lifetime imaging, luminescence imaging, vibrational absorption imaging or magnetic resonance imaging) have emerged as next-generation tools to unravel complex biological processes in space and time from particular vantage points. In contrast to genetically encodable fluorescent proteins, the probes for these advanced imaging modalities, however, generally lack biocompatible targeting strategies to specific biomolecules. Since proteins are the most diversified functional biomolecules, protein-specific targeting capability, if achievable, will tremendously enrich the applications of corresponding imaging methods.

Chemical tags have emerged as a new generation protein labeling strategy compatible with live cells. Chemical tags are composed of a defined polypeptide sequence that is fused to a protein of interest, and which can be subsequently modified with a chemical reagent, such as an appropriately derivatized fluorescent dye. The first chemical tag, FlAsH, was invented in 1998 by the Tsien lab[2]. Since then, several commonly used chemical tags have been developed, including self-labeling FlAsH/ReAsH[3], SNAP/CLIP tag[4,5], TMP-tag[6], HaloTag[7], β-lactamase tag[8] and enzyme-mediated labeling methods based on lipoic acid ligase[9]. Methodologically, chemists have used a variety of strategies to engineer and

optimize chemical tags, including directed evolution[10], proximity-induced reactivity[11] and pro-drug loading[9,12]. With efforts from many research groups, chemical tags have reached a relatively mature stage, and the question has shifted from 'How to label' to 'What to label with', as discussed in several recent review articles[13-15]. In our opinion, the most powerful feature of chemical tags, compared to the classic fluorescent proteins, is the rendered chemical diversity in the label/reporter moiety. We discuss in this chapter how this rendered chemical diversity perfectly matches this feature perfectly matches the demand of protein-specific imaging for a variety of advanced imaging methods.

In this chapter, we discuss the selected works that use chemical tags in combination with bio-imaging schemes beyond traditional fluorescence, such as wide-field or confocal microscopy. Reminiscent to the revolutionizing role of GFP to fluorescence microscopy, we highlight the bridging role of chemical tags that renders targeted protein specificity in modern advanced microscopies. And we also demonstrate the advantage of chemical tags in obtaining new and valuable information that would be difficult to collect otherwise (Table 1-1).

**Table 1-1. Implementations of chemical tags in advanced microscopy discussed in this chapter**

| Imaging Technique | Detected Signal | Representative Tag/Probe | Significance | Ref. |
|---|---|---|---|---|
| **PALM; STORM** | Precise location of single-molecule fluorophores | Halo/ azido DCDHF; TMP/Atto655 | Reconstructed super-resolution images | 21, 23, 26-29 |
| **Single-molecule tracking** | Precise single-molecule trajectories | SNAP/Dy547; Halo/Alexa488 | Dynamic information of target protein | 30-32 |
| **CosMos** | Colocalization of multi-color single-molecule fluorescence | TMP/Cy5; SNAP/DY549 | Mechanistic analysis of macromolecular machines | 33 |

| | | | | |
|---|---|---|---|---|
| **STED** | Peripherally depleted fluorescence | Halo/Atto655; SNAP/SiR | Super-resolution imaging | 29, 37-40 |
| **FLIM** | Fluorescence lifetime | TMP/Cy3 | Indicative of protein-specific micro-environment | 42 |
| **Triplet imaging** | Triplet-state lifetime | SNAP/TMR | Indicative of cellular oxygen consumption | 43 |
| **TR-FRET/LRET** | Time-resolved luminescence resonance energy transfer | SNAP/K (europium cryptate); TMP/Lumi4 (terbium chelate) | Background-free detection of FRET for protein-protein interaction | 44,45 |
| **OLID-FRET** | Modulated donor fluorescence due to acceptor photoswitching | SNAP/ NitroBIPS; SNAP/Cy3-NISO | Sensitive detection of FRET | 48,49 |
| **EM** | Electron beam | FlAsH-ReAsH | Protein-specific EM contrast | 3 |
| **Infrared near-field microscopy** | Infrared absorption | ACP/Alexa 488 | Photobleaching-free | 51 |
| **PET** | Gamma ray from positron emission | Halo/$^{64}$Cu-NOTA | *In Vivo* imaging of whole animal | 52 |
| **MRI** | Nuclear magnetic resonance | Halo/2CHTGd (gadolinium chelate) | Increased MRI sensitivity with protein-specificity | 53 |

### 1.3 Chemical tag-enabled imaging techniques based on single-molecule fluorescence

Single-molecule fluorescence imaging techniques have brought considerable excitement to biological research. These techniques enable characterization of biomolecules on the individual level, providing complementary data to that obtained from ensemble experiments. Because it only detects one molecule, the single-molecule fluorescence assay is technically demanding and requires high-photon-output fluorophores. A typical fluorescent protein molecule can emit roughly $4 \times 10^5$ photons before photobleaching[16], while the best organic dye molecules have a typical photon output on the order of $10^6$ to

$10^8$ [ref. 17]. Therefore, dye molecules conjugated with chemical tags provide high photon budgets, more precise localization, longer observation time and a higher signal-to-noise ratio. These merits make chemical tags excellent tools for the the study of proteins by single-molecule fluorescence.

Single-molecule fluorescence detection enables the reconstruction of sub-wavelength resolution images by two fundamentally similar approaches: PALM (photoactivation localization microscopy)[18,19] and STORM (stochastic optical reconstruction microscopy) (Figure 1-1a) [20]. In PALM, target proteins are labeled with photoactivatable fluorophores, which are then photo-activated sparsely and repeatedly, allowing the record of a collection of single-molecule resolved images. The fluorophores are finally localized to a precise location using software and the super-resolution image is generated. Compared to the photoactivatable fluorescent proteins, chemical tags allow more accurate localization due to the larger number of detected photons. From a chemical point of view, chemical tag-based labeling methods provide diverse photo-chemical strategies toward dye photoactivation. In 2010, the Moerner group demonstrated the first example of live bacteria PALM imaging of a labeled target protein using the chemical tag, HaloTag/azido DCDHF conjugate[21]. Azido DCDHF has an extraordinary high quantum yield of photoactivation under UV exposure[22]. Therefore, a low-intensity UV light source can be used, reducing the UV-induced damage to living cells. More recently, the Johnsson group utilized a caged rhodamine derivative as an alternative probe for PALM imaging in conjunction with the SNAP-tag[23]. Similar to PALM, STORM takes advantage of the reversible photoswitching of fluorescent dyes. It has recently been shown that photoswitching is a rather universal process for a wide spectrum of organic dyes, especially

rhodamines, cyanines and oxazines[24,25]. Live cell dSTORM (direct STORM) imaging of labeled intracellular protein H2B was demonstrated using a TMP–Atto655 conjugate, taking advantage of the photoswitching behavior observed in the presence of cellular oxygen and reductants[26]. Several alternative chemical tag/dye combinations have been successfully applied to live cell dSTORM[27,28], with a noteworthy example being the newly developed SNAP-tag/NIR fluorophore silicon–rhodamine[29]. With the growing availability of PALM/STORM microscope and chemical tag-dye conjugates, we expect super-resolution imaging to become a routine protocol for live cell studies in the near future.

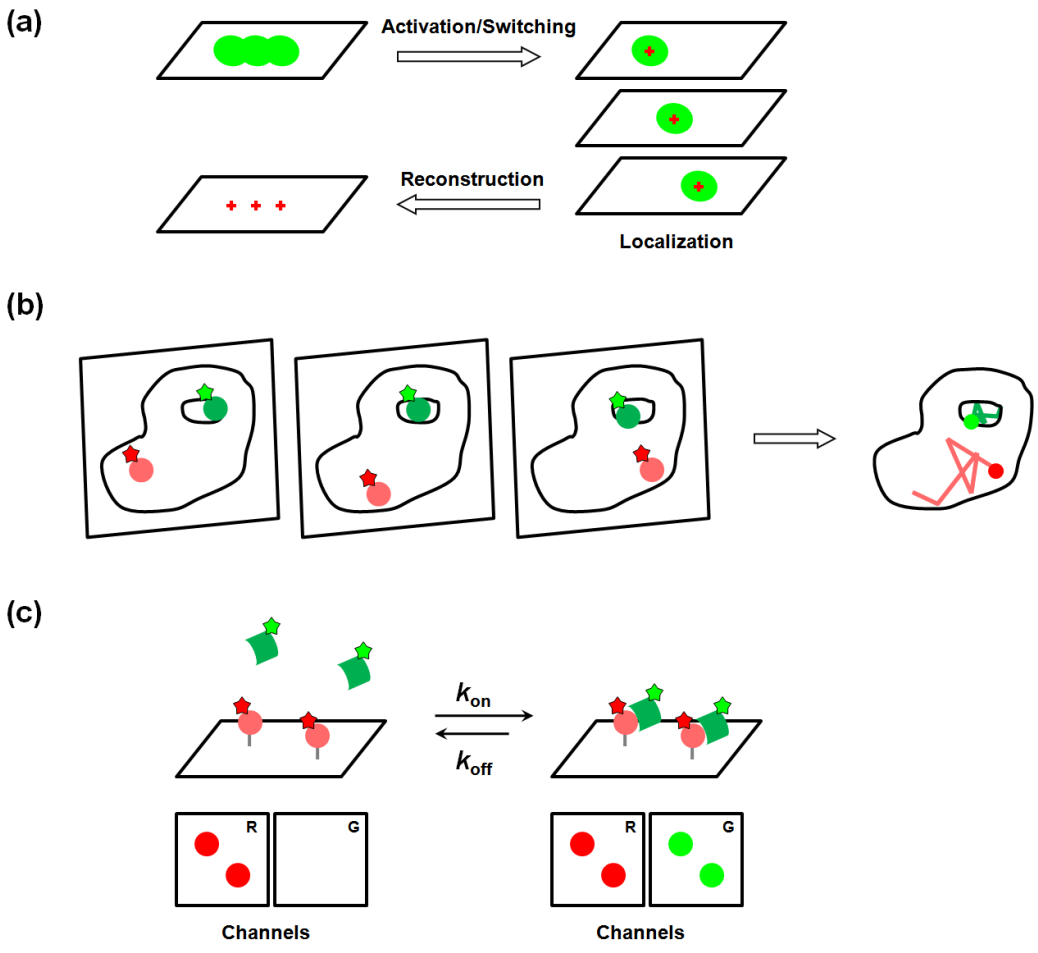



**Figure 1-1. Imaging strategies based on single-molecule fluorescence of labeled target proteins. (a)** PALM/STORM image is reconstructed based on precise localization of a stack of

images rendered by stochastic photoactivation/photoswitching on single-molecule level. **(b)** Single-molecule tracking is indicative of spatial‐temporal dynamics of different proteins in live cell. **(c)** Colocalization of multi-color single-molecule fluorescence reveals kinetic information of complex biochemical machineries.

The high-photon output of synthetic fluorophores could further enable prolonged sub-resolution tracking of single proteins with high temporal-resolution inside live cells (Figure 1-1b). Appelhans *et al*. observed single-molecule diffusion behavior of mitochondrial proteins using HaloTag-rhodamine labels[30]. Benke *et al*. later reported dual-color single-molecule tracking of cellular proteins using multiple chemical tags[31]. In pursuit of brighter and more photostable material, Liu *et al*. reported a targeting strategy of quantum dots that combines both the lipoic acid ligase and the HaloTag. This stepwise strategy enables the tracking of membrane targets at the single-molecule level and is considered a promising new method for tracking membrane-bound receptors in neurons[32].

The high photo-stability of organic dyes allows the observation of a single protein molecule over an extended period of time. The biochemical interactions within this time period can be seen under the same microscope if the interaction partners are labeled with other fluorescent markers with different colors, characterized as a colocalization event of multi-color single-molecule fluorescence (Figure 1-1c). This method, dubbed CoSMoS (colocalization single-molecule spectroscopy), is especially useful for the study of protein complexes which are difficult to reconstitute *in vitro*, as each component of the complex could be fluorescently labeled directly in cell extracts using chemical tags. Hoskins et al. have deciphered the dynamic assembly process of spliceosomes, which are mega-Dalton protein–RNA complexes for mRNA maturation, using CosMos with the TMP-tag and the SNAP-tag as labeling methods for individual protein components. This work highlights the

orthogonality between different chemical tags, enabling their simultaneous use for multi-color imaging[33]. In addition to probing biochemical interactions, single-molecule approaches are also useful for detecting low efficiency protein modifications due to its superior detection sensitivity. For example, Yang and Zhang demonstrated single-molecule measurements of simultaneous SUMOylation (small ubiquitin-like modifier) using the SNAP/CLIP tags, with a sensitivity ~100 fold greater than immunoblotting assays[34]. These examples are pioneering yet promising attempts to use chemical tags and single-molecule imaging to decipher protein functions and modifications.

## 1.4 Spatiotemporally resolved ensemble microscopy techniques

STED microscopy is the first demonstrated super-resolution technique for far-field fluorescence imaging. It uses a high-power doughnut-shaped stimulated emission beam to deplete the peripheral fluorescence of the focal point, therefore narrows the effective point spread function to improve spatial resolution[35,36]. In principle, stimulated emission depletion can be applied to all fluorophores (Figure 1-2a,b), but photostable dyes are preferable in practice since a high-intensity depletion beam is used. To date, live cell STED imaging has been demonstrated with several chemical tag/fluorophore conjugates[29,37-39]. Multi-color STED imaging with orthogonal chemical tags was also demonstrated[40].

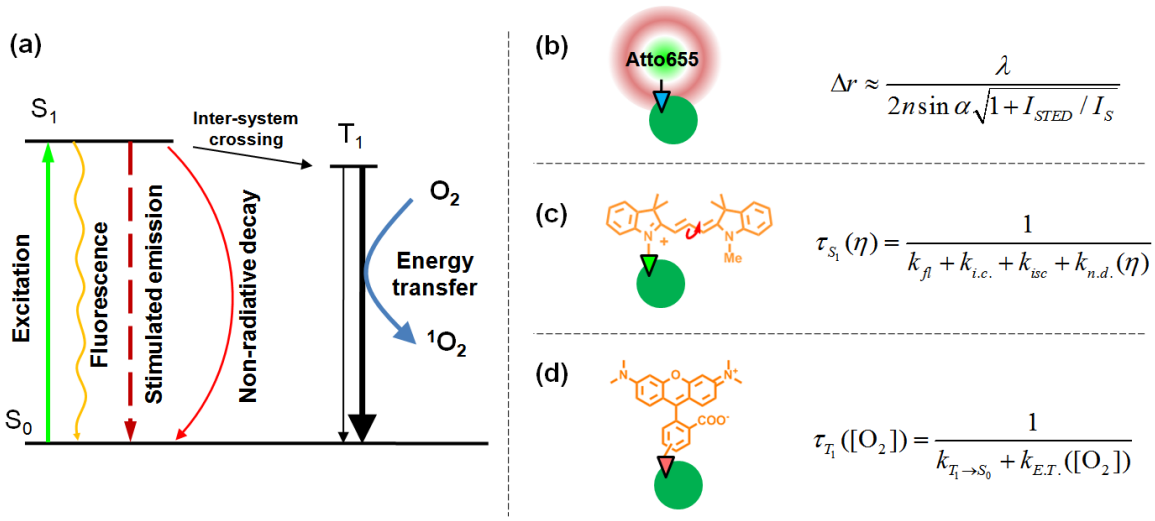



**Figure 1-2. Principles of selected spatiotemporally resolved imaging techniques. (a)** Energy diagram of a typical fluorophore. **(b)** HaloTag–Atto655 is used in STED microscopy to break the diffraction limit. Resolution of STED microscopy is a function of STED beam intensity. **(c)** TMP–Cy3 is used with FLIM microscopy to sense local micro-environment in live cells. Cy3 $S_1$ state lifetime is a function of local viscosity. **(d)** SNAP–TMR is able to sense oxygen concentration using triplet imaging microscopy. TMR $T_1$ state lifetime is a function of local oxygen concentration.

Besides implementations in super-resolution imaging, chemical tags have also advanced time-resolved fluorescence techniques. Fluorescence lifetime imaging microscopy (FLIM) is a powerful time-resolved fluorescence technique that characterizes the singlet excited state of fluorophores[41]. The fluorescence lifetime of the singlet excited state, defined as the time delay between the absorption of a photon and the emission event, is highly sensitive to the surrounding micro-environments. It can be measured in frequency-domain by modulating the excitation source sinusoidally at high frequency (>10 MHz) and recording the phase delay of the fluorescence signal. Gatzogiannis *et al*. have developed a protein micro-environment sensor based on fluorescence lifetime measurements of a TMP–Cy3 probe. Cy3 has a prolonged fluorescence lifetime in a high viscosity environment due to the lowered efficiency of its non-radiative isomerization pathway (Figure 1-2 a,c). Using TMP-

tag, Cy3 is targeted to the nucleus and cell membrane, and heterogeneity inside the cell nucleus is visualized using FLIM[42]. This work highlights the greater environmental accessibility of chemical tag-labeled organic dyes over barreled fluorescent proteins as biophysical probes. Using varying pulsed excitation, the triplet-state lifetime of fluorophores can also be measured by utilizing the fact that triplet-state build-up is a function of excitation pulse width. Geissbuehler *et al*. developed a wide-field imaging-fitting protocol to measure the triplet-state lifetime of tetramethylrhodamine (TMR) conjugated to a cytosol protein with SNAP-tag. Because oxygen induces triplet-state lifetime changes via an energy transfer process ( Figure 1-2a,d), this approach enables fast mapping of cellular oxygen concentration during muscle cell contraction[43].

Time-resolved fluorescence is especially useful in FRET-based methods to study protein–protein interactions (Figure 1-3a). Time-resolved FRET relies on a lanthanide-based donor which has a long-lived (on the order of ms) luminescence. Therefore, donor and acceptor emissions can be collected following pulsed excitations, minimizing the detection of cross-excitation of the acceptor and autofluorescence (Figure 1-3b). Maurel *et al*. designed a donor-acceptor pair, europium cryptate-d2, to GPCRs using SNAP-tag to study their interaction and oligomerization on the cell surface[44]. Rajapakse *et al*. have reported a TMP ‑ Lumi4 probe for studying cytosolic protein–protein interaction by luminescence resonance energy transfer (LRET) between a terbium complex and GFP[45]. Recently, TMP-lanthanide probes with improved cell permeability were reported[46]. With higher signal-to-noise ratio, time-resolved FRET is gaining increasing attention as a promising detection strategy of protein–protein interactions[47].

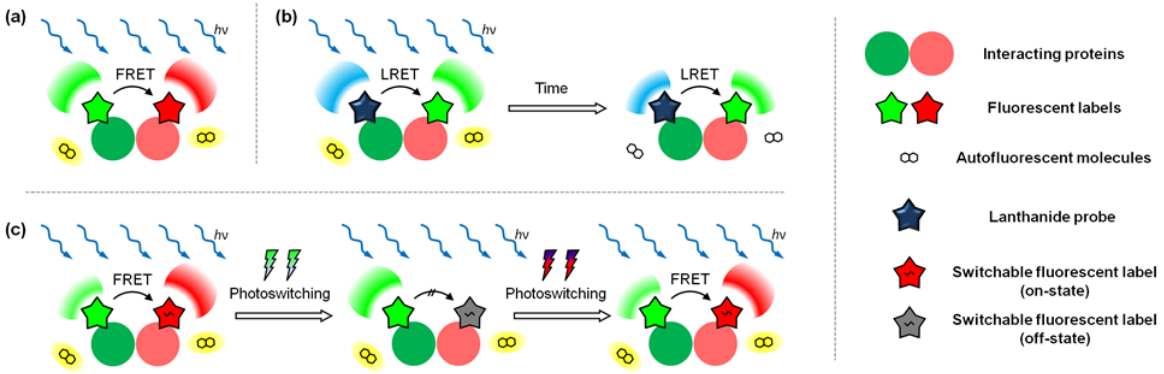



**Figure 1-3. Improving FRET using novel detection methods in combination with labeled fluorophores. (a)** Traditional FRET suffers from high background resulting from autofluorescence as well as spectrum bleed-through. **(b)** Using long-lived lanthanide probes as donors, emission can be detected shortly after the removal of the excitation light, minimizing the autofluorescence and cross-excitation. **(c)** Using OLID-FRET, the absorption spectrum of photoswitchable acceptor is directly modulated with light and FRET signal is detected from the modulated donor fluorescence.

Conceptually different from time-resolved FRET, OLID (optical lock-in detection)-FRET microscopy was developed as an alternative method to improve the detection accuracy of the FRET signal. In OLID-FRET, a photoswitchable fluorophore that can be reversibly photoswitched by light is used as the FRET acceptor. Donor fluorescence with and without sensitizing the acceptor was measured repeatedly in the same cell, allowing unambiguous resolution of the FRET signal (Figure 1-3c). Mao et al. reported a live cell OLID-FRET system using GFP as the donor and a photoswitchable NitroBIPS, conjugated to GFP via SNAP-tag, as the acceptor. This method enhances the sensitivity of FRET down to 1% FRET efficiency[48]. Recently Cy3/NISO was demonstrated to be a suitable OLID-FRET pair and can be used to label membrane proteins via the SNAP-tag[49].

## 1.5 Beyond fluorescence contrast

Although fluorescence is considered one of the most sensitive optical detection methods, it has several drawbacks including limited optical resolution, poor penetration

depth and inevitable fluorophore photobleaching. Development of non-fluorescence-based imaging methods could complement fluorescence imaging in these regards. However, often times there are few genetically encodable protein tags for these non-fluorescence methods. We aim to demonstrate the growing interest of using chemical tags in non-fluorescence-based imaging methods with targeted protein-specificity.

In pursuit of superior resolution beyond optical microscopy, electron microscopy (EM) was developed based on the fact that electron beams have orders of magnitude shorter wavelengths compared to that of visible light. By using the electron beam as the illumination source, EM can resolve sub-cellular structures down to 1 nm. The contrast of EM is usually rendered by staining with osmium tetroxide, which intrinsically lacks protein specificity. Gaietta *et al*. implemented ReAsH as a contrast reagent for electron microscopy by taking advantage of its photo-catalytic effect of diaminobenzidine oxidation toward an osmophilic polymeric product. Cooperation of fluorescence and electron microscopy of FlAsH/ReAsH labeled connexin43 revealed the transportation and incorporation processes of connexin43 into existing gap junctions[3]. This work is considered the benchmark for using chemical tags to achieve a resolution beyond optical microscopy.

As a non-bleaching alternative to traditional fluorescence approaches, infrared-based microscopy is being explored for potential imaging applications. Infrared (IR) absorption does not subject molecules to irreversible damage. The IR bands of biomolecules, however, are often super-imposed with each other and hard to distinguish. Generosi *et al*. observed that Alexa488 molecules had a specific IR absorption band which minimally overlapped with cellular IR absorption. Alexa488 was labeled to glutamate receptors on neurons with the ACP-tag[50] and live cell images were recorded using infrared scanning near-field

microscopy[51]. This work presents an interesting application of chemical tags towards photobleaching-free non-fluorescence optical imaging.

To achieve deeper penetration for live animal imaging, positron emission tomography (PET) is one of the most commonly used techniques. In PET, isotope probes are localized based on the emitted γ-rays, which are generated from the annihilation event between a positron and an electron. Hong *et al*. recently reported a HaloTag–$^{64}$Cu NOTA probe for PET imaging in live animals. Tumor cells expressing HaloTag protein could be detected in live mice using injected $^{64}$Cu NOTA probe[52]. Another widely used method for live animal imaging is magnetic resonance imaging (MRI), which uses penetrative magnetic field to magnetize and probe selected atomic nuclei and reconstruct images. Recently, a protein-targetable MRI contrast reagent based on HaloTag–gadolinium chelate (2CHTGd) was developed and characterized *in vitro*[53]. While further *in vivo* applications of the chemical tag-targeted MRI probe are still being evaluated, this work, along with the PET imaging approach discussed above, exemplify the potential of chemical tags in promoting protein-specific imaging in live animals.

## 1.6 Concluding Remarks

By introducing diverse reporting moieties specifically to their target proteins, chemical tags have unforeseeable potential in promoting novel techniques towards various biological problems. Stimulating to each other, the co-evolution of chemical tagging and imaging strategies is becoming a fruitful source of innovation for the toolbox of biological research.

## 1.7 References


1. Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509.
2. Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269.
3. Gaietta, G.; Deerinck, T. J.; Adams, S. R.; Bouwer, J.; Tour, O.; Laird, D. W.; Sosinsky, G. E.; Tsien, R. Y.; Ellisman, M. H. *Science* **2002**, *296*, 503.
4. Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol.* **2003**, *21*, 86.
5. Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K. *Chem. Biol.* **2008**, *15*, 128.
6. Miller, L. W.; Cai, Y.; Sheetz, M. P.; Cornish, V. W. *Nat. Methods* **2005**, *2*, 255.
7. Los, G. V; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. *ACS Chem. Biol.* **2008**, *3*, 373.
8. Mizukami, S.; Watanabe, S.; Hori, Y.; Kikuchi, K. *J. Am. Chem. Soc.* **2009**, *131*, 5016.
9. Mollwitz, B.; Brunk, E.; Schmitt, S.; Pojer, F.; Bannwarth, M.; Schiltz, M.; Rothlisberger, U.; Johnsson, K. *Biochemistry* **2012**, *51*, 986.
10. Chen, Z.; Jing, C.; Gallagher, S. S.; Sheetz, M. P.; Cornish, V. W. *J. Am. Chem. Soc.* **2012**, *134*, 13692.
11. Uttamapinant, C.; White, K. A.; Baruah, H.; Thompson, S.; Fernández-Suárez, M.; Puthenveetil, S.; Ting, A. Y. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 10914.
12. Watanabe, S.; Mizukami, S.; Akimoto, Y.; Hori, Y.; Kikuchi, K. *Chemistry* **2011**, *17*, 8342.
13. Jing, C.; Cornish, V. W. *Acc. Chem. Res.* **2011**, *44*, 784.
14. Hinner, M. J.; Johnsson, K. *Curr. Opin. Biotechnol.* **2010**, *21*, 766.
15. Fernández-Suárez, M.; Ting, A. Y. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 929.
16. Kubitscheck, U.; Kückmann, O.; Kues, T.; Peters, R. *Biophys. J.* **2000**, *78*, 2170.
17. Deschenes, L. A.; Vanden Bout, D. A. *Chem. Phys. Lett.* **2002**, *365*, 387.
18. Betzig, E.; Patterson, G. H.; Sougrat, R.; Lindwasser, O. W.; Olenych, S.; Bonifacino, J. S.; Davidson, M. W.; Lippincott-Schwartz, J.; Hess, H. F. *Science (80-. ).* **2006**, *313*, 1642.
19. Hess, S. T.; Girirajan, T. P. K.; Mason, M. D. *Biophys. J.* **2006**, *91*, 4258.
20. Rust, M. J.; Bates, M.; Zhuang, X. *Nat. Methods* **2006**, *3*, 793.
21. Lee, H. D.; Lord, S. J.; Iwanaga, S.; Zhan, K.; Xie, H.; Williams, J. C.; Wang, H.; Bowman, G. R.; Goley, E. D.; Shapiro, L.; Twieg, R. J.; Rao, J.; Moerner, W. E. *J. Am. Chem. Soc.* **2010**, *132*, 15099.

22. Lord, S. J.; Conley, N. R.; Lee, H. D.; Samuel, R.; Liu, N.; Twieg, R. J.; Moerner, W. E. *J. Am. Chem. Soc.* **2008**, *130*, 9204.

23. Banala, S.; Maurel, D.; Manley, S.; Johnsson, K. *ACS Chem. Biol.* **2012**, *7*, 289.

24. Heilemann, M.; van de Linde, S.; Mukherjee, A.; Sauer, M. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 6903.

25. Dempsey, G. T.; Vaughan, J. C.; Chen, K. H.; Bates, M.; Zhuang, X. *Nat. Methods* **2011**, *8*, 1027.

26. Wombacher, R.; Heidbreder, M.; van de Linde, S.; Sheetz, M. P.; Heilemann, M.; Cornish, V. W.; Sauer, M. *Nat. Methods* **2010**, *7*, 717.

27. Klein, T.; Löschberger, A.; Proppert, S.; Wolter, S.; van de Linde, S.; Sauer, M. *Nat. Methods* **2011**, *8*, 7.

28. Jones, S. A.; Shim, S.-H.; He, J.; Zhuang, X. *Nat. Methods* **2011**, *8*, 499.

29. Lukinavičius, G.; Umezawa, K.; Olivier, N.; Honigmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Corrêa, I. R.; Luo, Z.-G.; Schultz, C.; Lemke, E. a.; Heppenstall, P.; Eggeling, C.; Manley, S.; Johnsson, K. *Nat. Chem.* **2013**, 1.

30. Appelhans, T.; Richter, C. P.; Wilkens, V.; Hess, S. T.; Piehler, J.; Busch, K. B. *Nano Lett.* **2012**, *12*, 610.

31. Benke, A.; Olivier, N.; Gunzenhäuser, J.; Manley, S. *Nano Lett.* **2012**, *12*, 2619.

32. Liu, D. S.; Phipps, W. S.; Loh, K. H.; Howarth, M.; Ting, A. Y. *ACS Nano* **2012**, *6*, 11080.

33. Hoskins, A. A.; Friedman, L. J.; Gallagher, S. S.; Crawford, D. J.; Anderson, E. G.; Wombacher, R.; Ramirez, N.; Cornish, V. W.; Gelles, J.; Moore, M. J. *Science* **2011**, *331*, 1289.

34. Yang, Y.; Zhang, C. *Angew. Chemie* **2013**, *125*, 719.

35. Hell, S. W.; Wichmann, J. *Opt. Lett.* **1994**, *19*, 780.

36. Klar, T. A.; Jakobs, S.; Dyba, M.; Egner, A.; Hell, S. W. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8206.

37. Schröder, J.; Benink, H.; Dyba, M.; Los, G. V. *Biophys. J.* **2009**, *96*, L01.

38. Hein, B.; Willig, K. I.; Wurm, C. a; Westphal, V.; Jakobs, S.; Hell, S. W. *Biophys. J.* **2010**, *98*, 158.

39. Testa, I.; Wurm, C. A.; Medda, R.; Rothermel, E.; von Middendorf, C.; Fölling, J.; Jakobs, S.; Schönle, A.; Hell, S. W.; Eggeling, C. *Biophys. J.* **2010**, *99*, 2686.

40. Pellett, P. A.; Sun, X.; Gould, T. J.; Rothman, J. E.; Xu, M.-Q.; Corrêa, I. R.; Bewersdorf, J. *Biomed. Opt. Express* **2011**, *2*, 2364.

41. Berezin, M. Y.; Achilefu, S. *Chem. Rev.* **2010**, *110*, 2641.

42. Gatzogiannis, E.; Chen, Z.; Wei, L.; Wombacher, R.; Kao, Y.-T.; Yefremov, G.; Cornish, V. W.; Min, W. *Chem. Commun. (Camb).* **2012**, *48*, 8694.

43. Geissbuehler, M.; Spielmann, T.; Formey, A.; Märki, I.; Leutenegger, M.; Hinz, B.; Johnsson, K.; Van De Ville, D.; Lasser, T. *Biophys. J.* **2010**, *98*, 339.

44. Maurel, D.; Comps-Agrar, L.; Brock, C.; Rives, M.-L.; Bourrier, E.; Ayoub, M. A.; Bazin, H.; Tinel, N.; Durroux, T.; Prézeau, L.; Trinquet, E.; Pin, J.-P. *Nat. Methods* **2008**, *5*, 561.

45. Rajapakse, H. E.; Gahlaut, N.; Mohandessi, S.; Yu, D.; Turner, J. R.; Miller, L. W. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 13582.

46. Mohandessi, S.; Rajendran, M.; Magda, D.; Miller, L. W. *Chemistry* **2012**, 1.

47. Lowder, M. a; Appelbaum, J. S.; Hobert, E. M.; Schepartz, A. *Curr. Opin. Chem. Biol.* **2011**, *15*, 781.

48. Mao, S.; Benninger, R. K. P.; Yan, Y.; Petchprayoon, C.; Jackson, D.; Easley, C. J.; Piston, D. W.; Marriott, G. *Biophys. J.* **2008**, *94*, 4515.

49. Yan, Y.; Petchprayoon, C.; Mao, S.; Marriott, G. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2013**, *368*, 20120031.

50. George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K. *J. Am. Chem. Soc.* **2004**, *126*, 8896.

51. Generosi, J.; Margaritondo, G.; Kropf, M.; Hirling, H.; Catsicas, S.; Johnsson, K.; Tolk, N. H.; Piston, D. W.; Cricenti, a. *J. Appl. Phys.* **2008**, *104*, 106102.

52. Hong, H.; Benink, H. A.; Zhang, Y.; Yang, Y.; Uyeda, H. T.; Engle, J. W.; Severin, W.; Mcdougall, M. G.; Barnhart, T. E.; Klaubert, D. H.; Nickles, R. J.; Fan, F. **2011**, *3*, 392.

53. Strauch, R. C.; Mastarone, D. J.; Sukerkar, P. a; Song, Y.; Ipsaro, J. J.; Meade, T. J. *J. Am. Chem. Soc.* **2011**, *133*, 16346.

# Chapter 2

## Second-Generation Covalent TMP-Tag for Live Cell Imaging

The content of this chapter has been published in:

Chen, Z.; Jing, C.; Gallagher, S. S.; Sheetz, M. P.; Cornish, V. W. “Second-Generation Covalent TMP-Tag for Live Cell Imaging”. *J. Am. Chem. Soc.* **2012**, 134, 13692.

## 2.1 Chapter outlook

Chemical tags are now viable alternatives to fluorescent proteins for labeling proteins in living cells with organic fluorophores that have improved brightness and other specialized properties. Recently, we successfully rendered our TMP-tag covalent with a proximity-induced reaction between the protein tag and the ligand-fluorophore label. This initial design, however, suffered from slow *in vitro* labeling kinetics and limited live cell protein labeling. Thus, here we report a second-generation covalent TMP-tag that has a fast labeling half-life and can readily label a variety of intracellular proteins in living cells. Specifically, we designed an acrylamide-trimethoprim-fluorophore (A-TMP-fluorophore v2.0) electrophile with an optimized linker for fast reaction with a cysteine (Cys) nucleophile engineered just outside the TMP-binding pocket of *E. coli* dihydrofolate reductase (eDHFR) and developed an efficient chemical synthesis for routine production of a variety of A-TMP-probe v2.0 labels. We then screened a panel of eDHFR:Cys variants and identified eDHFR:L28C as having an 8 minute half-life for reaction with A-TMP-biotin v2.0 *in vitro*. Finally, we demonstrated live cell imaging with A-TMP-fluorescein, A-TMP-Dapoxyl and A-TMP-Atto655 of various cellular protein targets. With its robustness, this second-generation covalent TMP-tag adds to the limited number of chemical tags that can be used to covalently label intracellular proteins efficiently in living cells. Moreover, the success of this second-generation design further validates proximity-induced reactivity and organic chemistry as tools not only for chemical tag engineering but also more broadly for synthetic biology.

I have been the main contributor of the project. Chaoran Jing and I designed the second generation protein/A-TMP pair based on the work of Sarah Gallagher, Michael Sheetz and Virginia Cornish. Chaoran Jing performed the subclonings and purified the

proteins for *in vitro* assays. I designed and synthesized the A-TMP-probes, performed the labeling and imaging experiments. I and Virginia Cornish wrote the manuscript with contributions from all co-authors.

## 2.2 Introduction

Chemical tags are emerging from the proof-of-principle stage to viable reagents for labeling proteins in living cells with fluorophores with high photon outputs and other specialized properties[1-3] to enable experiments difficult or not possible with the fluorescent proteins (FPs) [4]. With chemical tags, rather than tagging the target protein with an FP, the target protein is tagged with a polypeptide, which is subsequently labeled with a cell-permeable fluorophore ligand or substrate. Thus, chemical tags combine the advantage of specificity through genetic encoding with a modular organic fluorophore. Chemical tags now in use include the seminal peptide chelator-based FlAsH/ReAsH system[5], the enzyme suicide substrate-based SNAP/CLIP-tags[6,7] and Halo-tag[8], the small molecule inhibitor-based TMP-tag[9], and the enzyme-mediated polypeptide labeling-based lipoic acid ligase tag[10]. Exciting recent applications of the chemical tags include single-molecule imaging of spliceosome function in yeast cell extracts[11], magnetically modulating mammalian cells using decorated iron oxide nanoparticles[12], imaging LDL receptor oligomerization during endocytosis[13] and super-resolution imaging of cellular proteins [14]. While new chemical tags are regularly being introduced in the literature[1-3], our TMP-tag still stands out as one of the few chemical tags able to label intracellular, as opposed to cell-surface, proteins with high selectivity.

With TMP-tag the target protein is tagged with *E. coli* dihydrofolate reductase (eDHFR) through standard genetic encoding and then labeled by binding to a cell-permeable trimethorpim-fluorophore (TMP-fluorophore) conjugate. eDHFR is an attractive protein tag because it is 18 kD (about two-thirds the size of GFP) and monomeric, thus minimally disrupts biological function of the tagged protein and pathway. With low nM affinities for

eDHFR, the TMP-fluorophore conjugates can be used at near stoichimetric quantities to efficiently label tagged cellular proteins, which average a ~1 μM concentration in the cell. At the same time, with > 1,000-fold selectivity for *E. coli* over mammalian DHFRs, TMP-tag shows minimal background labeling of endogenous proteins and no apparent cellular toxicity in mammalian cell lines. TMP is commercially available and can be readily modified without disrupting binding to eDHFR, facilitating the preparation of a wide variety of TMP analogs[15]. Finally, there is a wealth of biochemical and structural knowledge of the interaction between TMP and eDHFR, which facilitates further engineering of the TMP-tag[16]. Fine tuning of the fluorophore hydrophobicity and linker structure have produced optimized versions of TMP-tag for lower unspecific background staining and better cell permeability[17]. On the strength of its robustness, TMP-tag labels have been developed to enable super-resolution imaging[14] and two-photon imaging of cellular proteins[18], chromophore-assisted laser inactivation (CALI) of components of the focal adhesion complex[19], and single-molecule imaging of spliceosome assembly[11].

To provide a more permanent label for advanced applications such as single-molecule tracking or pulse-chase labeling, we recently reported a covalent variant of the TMP-tag based on a proximity-induced reaction between the eDHFR tag and the TMP-fluorophore label[20]. Briefly, a unique Cys nucleophile was engineered just outside the TMP-binding pocket of eDHFR (eDHFR:L28C) in position to react with an acrylamide electrophile installed on the TMP-fluorophore label (acrylamide-TMP-fluorophore, or A-TMP-fluorophore). This design was based on the long-standing use of proximity-induced reactivity[21, 22] for the design of covalent inhibitors[23, 24] and more recent application to chemical biology tools[25]. Work of Belshaw and co-workers[26, 27] led us to believe that the

acrylamide electrophile would have the right balance in reactivity, being a sufficiently mild electrophile to minimize non-specific, background labeling of cellular components, but being reactive enough to undergo a rapid Michael addition upon TMP binding to eDHFR. Our initial design was successful, and we demonstrated that A-TMP-biotin reacted with eDHFR:L28C with a half-life of ~50 min *in vitro* and that A-TMP-fluorophore could covalently label a nuclear-localized eDHFR fusion protein in live cells with minimal background labeling of other cellular proteins. However, this first-generation covalent TMP-tag was unable to label cytoplasmic proteins tagged with eDHFR:L28C, limiting its utility. We hypothesized that this limited reactivity resulted from the slow half-life with which A-TMP reacted with eDHFR:L28C.

Thus, we sought to design a second-generation covalent TMP-tag with a rapid labeling half-life that would improve its utility for live cell imaging. Previous reports in the literature have shown that the half-life of both covalent inhibitors[24] and chemical biology tools[26] can be improved to a few minutes with optimization of the reaction geometry between the protein nucleophile and the organic electrophile. We present the design and synthesis of an optimized v2.0 A-TMP-probe in conjunction with the rational design and screening of a panel of eDHFR:Cys variants to generate a v2.0 covalent TMP-tag with a rapid reaction half-life. Finally, we challenge the robustness of this v2.0 covalent TMP-tag with live-cell, confocal fluorescence imaging of multiple intracellular proteins in different mammalian cell lines.

## 2.3 Results

### 2.3.1 Design of the Second-Generation Covalent TMP-Tag

Based on the success of our initial covalent TMP-tag, a second-generation covalent TMP-tag was also built around the Cys nucleophile and acrylamide electrophile (Figure 2-1), while optimizing the positioning of the nucleophile and electrophile to improve the reaction half-life. The Cys nucleophile and acrylamide electrophile had exceeded our expectations for minimal background labeling of endogenous cellular components and minimal cellular toxicity and yet still were able to undergo a fairly rapid binding-induced Michael addition. Previous literature on the design of covalent inhibitors and chemical biology reagents suggested that we could achieve a reaction half-life of a few minutes simply by optimizing the positioning of the Cys side-chain and acrylamide group undergoing the Michael addition[26]. For our initial covalent TMP-tag[20], we chose a conservative design containing a 21-atom linker between the 4'-OH group of TMP and the reactive β-carbon of the acrylamide functional group to ensure that the acrylamide would be available to react with the Cys nucleophile installed on the surface of eDHFR. Because of this long linker length, it was not surprising that the initial covalent TMP-tag had an *in vitro* labeling reaction half-life of almost 1 hour.

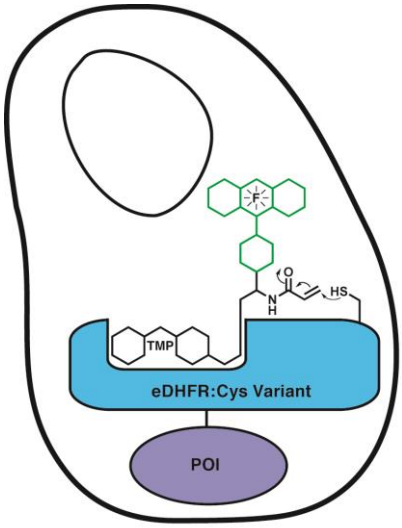



**Figure 2-1. Schematic representation of the covalent TMP-tag design**. Previously, we demonstrated that the non-covalent TMP-tag, which exploits the high-affinity, selective interaction between trimethoprim (TMP) and *E. coli* dihydrofolate reductase (eDHFR), could be rendered covalent by proximity-induced reaction between a Cys residue engineered on the eDHFR surface and a mild acrylamide electrophile installed on the TMP-fluorophore probe. Here, by optimizing the positioning of the Cys nucelophile and the acrylamide electrophile, we achieve rapid covalent labeling of the eDHFR tag by the TMP-fluorophore probe, rendering the covalent acrylamide TMP-tag (A-TMP-tag) a robust reagent for live cell imaging.

We chose to use rational design in combination with screening of a small number of variants to create a covalent TMP-tag with the minimum necessary distance between the Cys nucleophile and the acrylamide electrophile to achieve the desired reduction in reaction half-life. First, molecular modeling was applied to the high-resolution structure of eDHFR[16, 28] to identify residues that had solvent-accessible side chains in which the side chain faced the binding pocket to ensure the engineered Cys residue would be accessible to react with the acrylamide electrophile. A model of TMP bound to eDHFR was created by structurally aligning a high-resolution structure of *E. coli* DHFR[28] to a high-resolution structure of TMP bound to *L. casei* DHFR[29]. Only residues in close proximity to the binding pocket were selected, since precedent has shown that the closer the residue to the binding pocket, the

more rapid the rate of alkylation[26]. Four residues were selected that met this criterion: Ala19, Asn23, Leu28 and Arg52. An approximation was made of the minimum linker length between TMP and the electrophile that would allow proximity-induced covalent labeling to occur upon binding to a mutant eDHFR containing each of these Cys mutants (Table 2-S1). According to the model described above, we envisioned that a 10-atom spacer between the 4'-OH group of TMP and the β-carbon of the acrylamide would enable the electrophile to reach all the four engineered Cys nucleophiles (Figure 2-2).

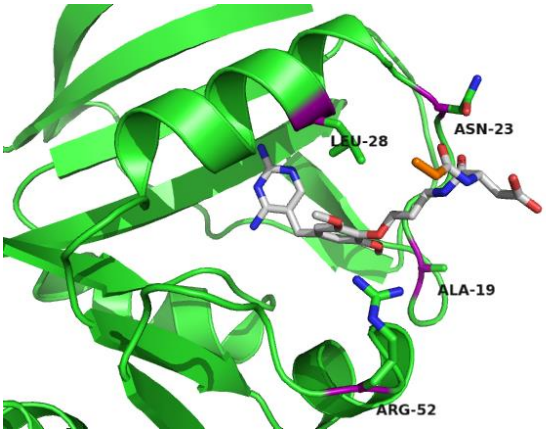


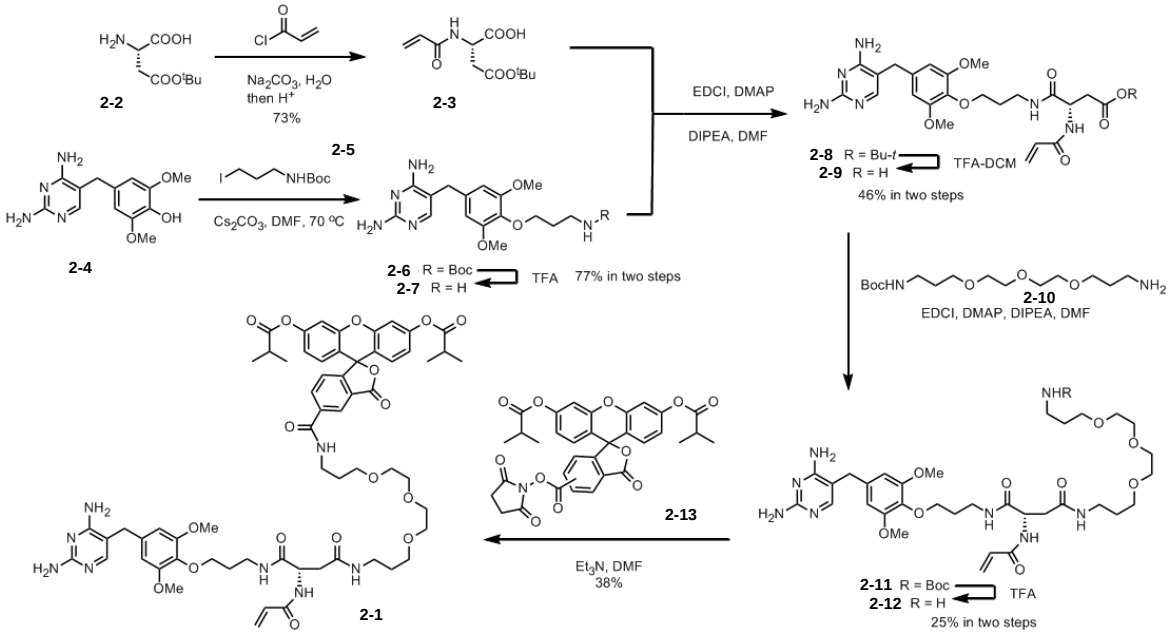

**Figure 2-2. Design of the optimized, second-generation covalent A-TMP-tag.** The acrylamide elecrophile on the TMP-fluorophore probe and the Cys nucleophile on the eDHFR surface were redesigned to bring the two in close proximity to achieve a rapid reaction half-life. Depicted is a cartoon of a second-generation A-TMP molecule (stick representation, electrophile highlighted in orange) with a 10-atom spacer between the TMP ligand bound in the active site of eDHFR (green ribbon diagram) and the acrylamide electrophile with the four residues chosen for mutation to Cys highlighted (stick representation, $\alpha$ carbon highlighted in purple). Since there is no reported high-resolution structure of TMP bound to eDHFR, this model was created by structurally aligning a high-resolution structure of *E. coli* DHFR[28] to a high-resolution structure of TMP bound to *L. casei* DHFR[29]. The acrylamide-TMP structure was built in Maestro[30] and then superimposed on TMP in the eDHFR model. The graphic was prepared using PyMOL[31].

### 2.3.2 Synthesis of the A-TMP-probe heterotrimer

Guided by molecular modeling, we designed an A-TMP-probe v2.0 heterotrimer with a 10-atom spacer between the 4'-OH group of TMP and the β-carbon of the acrylamide. We chose aspartic acid as the trifunctional core because amino acid derivatives could provide a convenient protection strategy for the sequential addition of the three different groups. In the proof-of-principle demonstration we chose fluorescein diisobutyrate as the probe, as cell behavior of fluorescein derivatives have been well-studied[17, 32]. We chose fluorescein diisobutyrate conjugate due to its higher chemical stability over fluorescein diacetate conjugate as DMF stock solution. Addition of the probe in the final step could be beneficial to the generality of the tag, facilitating the preparation of a variety of A-TMP derivatives with different probe molecules.

The synthetic route of the target molecule is summarized in Scheme 2-1. H-Asp(OBu-$^t$)-OH (**2-2**), a commercially available aspartic acid derivative, was treated with acryloyl chloride to yield carboxylic acid **2-3**. Amine **2-7** was prepared by *O*-alkylation of TMP phenol **2-4**[20] with a three-carbon Boc-amino iodide (**2-5**) followed by TFA deprotection of the Boc group. Coupling of carboxylic acid **2-3** and amine **2-7** with EDCI led to *tert*-butyl ester **2-8**, which was subjected to TFA deprotection to yield carboxylic acid **2-9**, a key intermediate towards the A-TMP-probe heterotrimer. A PEG linker was incorporated to carboxylic acid **2-9** by EDCI mediated coupling with a mono-protected PEG bis-amine (**2-10**) followed by TFA deprotection of the Boc group. The product, amine **2-12**, was coupled with protected fluorescein NHS-ester (**2-13**) and purified by HPLC to yield the final heterotrimer, compound **2-1,** in pure form. The heterotrimer was prepared from aspartic acid derivative **2-2** in 3.4% overall yield with the longest linear sequence consisting of seven

steps. This modular synthetic plan would allow us to prepare a variety of v2.0 A-TMP-Probe molecules. For an illustrative example, A-TMP-Biotin v2.0 (**2-S2**) was also prepared by a similar synthetic plan (Figure 2-S2).




**Scheme 2-1. Synthesis of optimized acrylamide-TMP-fluorescein heterotrimer (A-TMP-fluorescein v2.0, compound 2-1).**

### 2.3.3 *In vitro* screening of A-TMP v2.0 with eDHFR:Cys variants

After we obtained the A-TMP-Probe v2.0 molecule, we moved to *in vitro* labeling studies to determine the best eDHFR:Cys variant to pair with A-TMP (Figure 2-3A). We found that the most rapid reaction occurred between eDHFR:L28C and A-TMP-Biotin v2.0 among the tested eDHFR variants. Under the tested labeling conditions, the half-life of the labeling reaction was determined to be 8 minutes in the presence of NADPH.

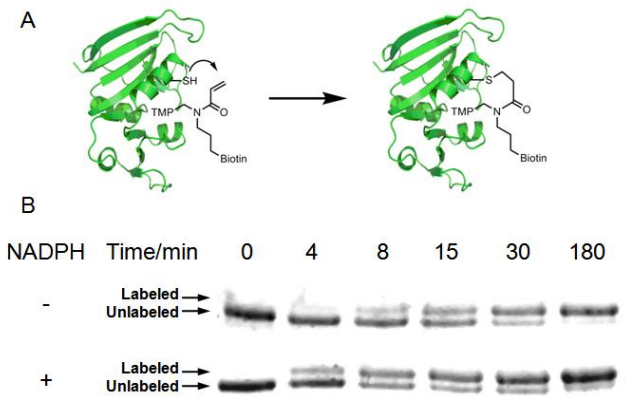



**Figure 2-3. Determination of the rate of covalent labeling between A-TMP-biotin v2.0 and eDHFR:L28C *in vitro*.** (A) Illustrative reaction scheme of the proximity-induced Michael addition of the thiol nucleophile of L28C to the acrylamide electrophile of A-TMP-biotin v2.0. (B) Analysis of the labeling reaction between A-TMP-biotin v2.0 and eDHFR:L28C by SDS-PAGE. Purified eDHFR:L28C (5 μM) was incubated with A-TMP-biotin v2.0 (10 μM) in PBS buffer (pH 7.4) with reduced glutathione (1 mM) at 37°C, with or without NADPH (50 μM). At proper time points, aliquots (20 μL) were removed from the reaction mixture, quenched with 6x SDS and boiled for 5 min. The time point aliquots were then analyzed by SDS-PAGE and Coomasie staining. Conveniently, covalent modification of eDHFR:L28C gave rise to a gel shift such that the reaction progress could be readily measured by densitometry analysis of Coomassie stained gels using Image-J. The labeling half-life was determined by linear regression, applying the pseudo first order model onto the ratio of the substrate and product. eDHFR:L28C was found to react with A-TMP-biotin v2.0 with a half life of 8 min at these physiologically relevant conditions.

Evaluation of the reactions between A-TMP V2.0 and eDHFR variants was conducted using purified proteins. The four designed eDHFR:Cys variants eDHFR:A19C, eDHFR:N23C, eDHFR:L28C and eDHFR:R52C were overexpressed via the T7 promoter using the corresponding *E. coli* expression vectors and purified using Ni-NTA spin columns. The proteins were judged to be more than 95% pure according to Coomassie staining of SDS-PAGE gels. The endogenous cysteines of eDHFR, Cys85 and Cys152, were mutated to serines in these vectors to minimize possible cross reactivity with the engineered Cys nucleophile. A-TMP-Biotin v2.0 was used as the tag in the conditional screening due to its better solubility in PBS buffer compared with the hydrophobic protected A-TMP-Fl.

Next, the *in vitro* labeling kinetics were determined using a SDS-PAGE gel shift assay to identify the fastest eDHFR/A-TMP pair (Figure 2-3B). Following the conditions reported in our previous study[20], 10 μM A-TMP-biotin was reacted with 5 μM eDHFR:Cys variant in PBS buffer, and the reaction mixture was quenched at appropriate intervals with 6X SDS and then subjected to SDS-PAGE. Gel shifts were produced due to covalent modification of the eDHFR:Cys varients, simplifying analysis of the reaction progression as described in Figure 2-3. We first tested the reaction between eDHFR:L28C and A-TMP-biotin v2.0. The labeling reaction was near quantitative after 3h, and the time required for 50% labeling was determined to be 17 min. We then tested the effect of NADPH to this labeling reaction, as NADPH is a native cofactor of eDHFR. In the presence of 50 μM NADPH, estimated to be the cellular concentration of NADPH[33,34], the reaction between eDHFR:L28C and A-TMP-biotin v2.0 was accelerated with a half-life of 8 min. Further screening found that all of the eDHFR variants (Figure 2-S3) reacted with A-TMP-biotin, but with significantly different reaction half-lives (Table 2-1). Intriguingly, NADPH promoted the reaction of A-TMP-biotin v2.0 with eDHFR:L28C, N23C and A19C, but slowed the reaction with eDHFR:R52C, perhaps indicative of complex conformational effects of NADPH binding. Overall, eDHFR:L28C was chosen as the fastest target for the designed second-generation A-TMP-probe molecule, especially in the presence of NADPH. This system would be particularly suitable for intracellular targets because of the abundance of NADPH in the reducing cellular environment.

**Table 2-1. Reaction half-lives [min] of A-TMP-biotin v2.0 with eDHFR:Cys variants.** Purified eDHFR:Cys variant was labeled under the same conditions as in Figure 2-3B.

| eDHFR Variants | | L28C | N23C | A19C | R52C |
|---|---|---|---|---|---|
| NADPH | - | 17 | 35 | 330 | 130 |
| | + | 8 | 20 | 100 | 220 |

### 2.3.4 Protein labeling in live cells with the second-generation covalent TMP-tag

Encouraged by the rapid *in vitro* labeling reaction between eDHFR:L28C and A-TMP-probe v2.0, the selected pair was next evaluated by labeling of cellular proteins. eDHFR:L28C was genetically fused to the C-termini of four different target proteins: histone H2B, Tomm20, $\alpha$-actinin and myosin light chain (MLC). Mammalian cell lines expressing the fusion proteins were successfully labeled with A-TMP-Fluorescein v2.0 in 10 min and were characterized by both microscopy and in-gel fluorescence analysis. These results demonstrate the generality of the v2.0 covalent TMP-tag for live cell imaging.

At first, we aimed for labeling of an abundant cellular target. We chose histone H2B, an essential nuclear protein which has been intensively investigated, as the first target. HEK 293T cells transiently transfected with plasmids encoding H2B-eDHFR:L28C fusion protein were incubated with 1 μM A-TMP-Fl v2.0 for 10 min. After staining, cells were washed twice with media and imaged by confocal microscopy. Distinct nucleic distribution of fluorescence was observed in transfected cells, with chromosomal patterns observed in a number of cells (Figure 2-4A). No significant background cytosol staining was detected. These observations indicated that A-TMP-Fl v2.0 was able to selectively bind to H2B-eDHFR:L28C fusion proteins with rapid kinetics in live cells.

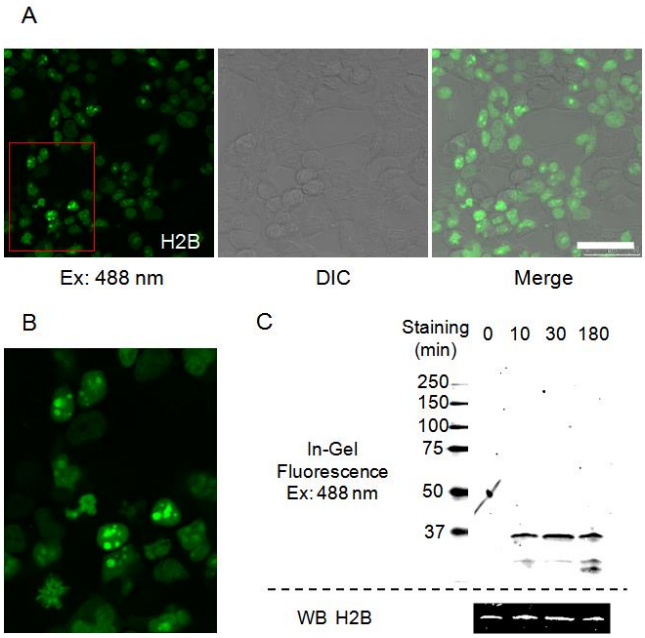



**Figure 2-4. Labeling of H2B with the covalent A-TMP-fluorescein v2.0 in HEK293T cells.** (A) Microscopic evidence of successful labeling of H2B tagged with eDHFR:L28C (H2B-eDHFR:L28C) by A-TMP-fluorescein v2.0 in Human Embryonic Kidney (HEK) 293T cells. HEK293T cells transiently transfected with a vector encoding the H2B-eDHFR:L28C fusion protein were incubated with 1 μM A-TMP-fluorescein v2.0 in the appropriate media for 10 min, washed twice, and then directly imaged using confocal fluorescence and differential interference contrast (DIC) microscopy. Fluorescein was excited at 488 nm. Scale bars, 50 μm. (B) Zoom-in view of the fluorescence image shown in (A). (C) In-gel fluorescence and Western blot analysis of the labeling reaction. Labeled HEK 293T cells were lysed and then analyzed by SDS-PAGE and in-gel fluorescence scanning with an excitation laser at 488 nm. Together, these results provide evidence that A-TMP-fluorescein v2.0 labels eDHFR:L28C tagged H2B rapidly, selectively, and covalently.

To confirm that the labeling reaction was covalent, HEK 293T cells expressing the H2B-eDHFR:L28C fusion protein were treated with 1 μM A-TMP-Fl v2.0 for 10 min, 30 min or 3 hr. After staining, cells were lysed and analyzed by SDS-PAGE and in-gel fluorescence (Figure 2-4B). A major band with a green channel emission was detected as the expected 35 kD H2B-eDHFR:L28C-A-TMP-Fl v2.0 conjugate, while no detectable background binding was observed in non-transfected cells (Figure 2-S4). The labeling products were further confirmed by western blot analysis using Anti-H2B antibody. Several

minor bands with lower molecular weight were also detected in cells expressing the H2B-eDHFR:L28C fusion protein, particularly in longer incubation. These bands are likely to be the degradation products of labeled H2B, as the control experiment showed undetectable background staining of endogenous proteins.

Our next goal was to test the versatility of the second-generation covalent TMP-tag for labeling diffuse cellular protein targets. Tomm20, a mitochondrial localized protein, was chosen as an organelle target in HEK 293T cells. Myosin light chain (MLC) and $\alpha$-actinin, two cytoplasmic proteins, was chosen as cell skeleton targets in fibroblast cells. Cells expressing eDHFR:L28C fusions were treated with 1 μM A-TMP-Fl v2.0 for 10 min and then examined using a confocal microscope. In each scenario, H2B-mCherry fusion protein was co-transfected. The fluorescence images indicate that all three eDHFR:L28C fusion proteins could be successfully labeled (Figure 2-5A). To further characterize the labeling specificity, cells expressing the different eDHFR:L28C fusions were lysed after 10 min treatment of 1 μM A-TMP-Fl v2.0. The lysates were subjected to SDS-PAGE and analyzed by in-gel fluorescence with a 488 nm laser. In all cases, a single main fluorescent band of the expected molecular weight was observed (Figure 2-5B), indicating a rapid labeling with high specificity among all the examined cellular protein targets in different mammalian cell lines.

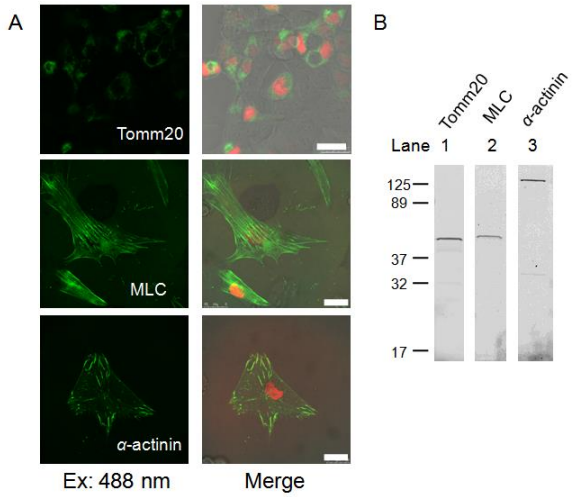



**Figure 2-5. Labeling of diffused protein targets with covalent A-TMP-Fl v2.0.** Three different proteins, Tomm20, MLC and $\alpha$-actinin, were successful labeled and imaged in two different mammalian cell lines. (A) Microscopic imaging of A-TMP-tag v2.0 labeling. HEK293T cells (for Tomm20) or Mouse Embryonic Fibroblast (MEF) cells (for MLC and $\alpha$-actinin) transiently co-transfected with vectors encoding POI-eDHFR:L28C and H2B-mCherry fusion proteins, respectively, were incubated with 1 μM A-TMP-fluorescein v2.0 in media for 10 min, washed twice with media and directly imaged using confocal and differential interference contrast (DIC) microscopy. Fluorescein was excited at 488 nm, mCherry was excited at 594 nm. Scale bars, 25 μm. (B) In-gel fluorescence analysis of A-TMP-Fl v2.0 labeling. The cells transfected with corresponding POI-eDHFR:L28C vectors were harvested after 10 min incubation with 1 μM A-TMP-fluorescein v2.0, lysed and analyzed by SDS-PAGE and in-gel fluorescence scanning with an excitation laser at 488 nm. These results show the target versatility of the second-generation covalent TMP-tag for live cell protein labeling.

With A-TMP-Fl v2.0 in hand, we performed labeling experiments to evaluate the performance of this covalent TMP-tag compared to non-covalent TMP-tag (Figure 2-6). HEK 293T Cells transiently expressing Tomm20-eDHFR and Tomm20-eDHFR:L28C were labeled with TMP-Fl[17] and A-TMP-Fl v2.0 for 10 min, respectively. In both experiments H2B-mCherry were co-transfected as counter stains. After washing with fresh media, cells were imaged with a confocal microscope. While non-covalent TMP-tag exhibits a similar labeling specificity in live cell imaging experiments with the covalent TMP-tag, the labeling pattern could not be distinguished 12 h after paraformaldehyde fixation. In contrast, the

labeling of Tomm20-eDHFR:L28C with A-TMP-Fl v2.0 withstands the fixation protocol, which potentially facilitates the applications of novel microscopic studies requiring long acquisition time.

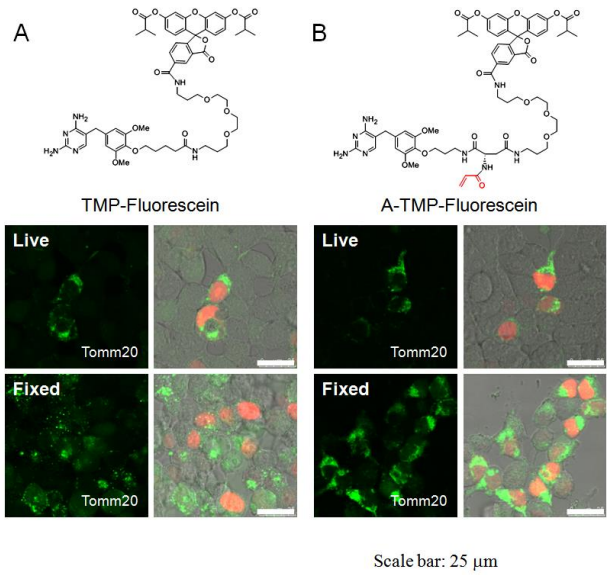



**Figure 2-6. Comparative studies of labeling with non-covalent TMP-tag and covalent TMP-tag.** (A) HEK 293T cells expressing Tomm20-eDHFR was labeled with TMP-Fl as in Figure 2-5A. After labeling, cells were imaged using confocal microscope (Live panel) or fixed with 4% paraformaldehyde in PBS for 10 min followed by washing with PBS for 12 hr before imaged (Fixed panel). (B) HEK 293T cells expressing Tomm20-eDHFR:L28C was labeled with A-TMP-Fl and examined with and without fixation treatment as in (A). Scale bars, 25μm.

Finally, we prepared A-TMP-Atto 655 and A-TMP-Dapoxyl to demonstrate the adaptability of second generation covalent TMP-tag over novel fluorophores for potential advanced imaging applications. Atto 655 has been demonstrated as an ideal organic fluorophore for live-cell super-resolution imaging due to its unique cellular-environment-compatible photoswitching mechanism[14]. After 3 h incubation of 1 μM A-TMP-Atto 655 with HEK 293T cells transiently expressing H2B-eDHFR:L28C or plasma membrane targeted eDHFR:L28C (PMLS-eDHFR:L28C)[9], selective labeling could be observed in both cases with confocal microscopy (Figure 2-7A). Dapoxyl dye, since its invention[35], has been

gaining growing attention due to its large and environmentally sensitive Stokes shift[36]. A-TMP-Dapoxyl was tested in labeling experiments with H2B-eDHFR:L28C as well as plasma membrane targeted eDHFR:L28C. Organelle-specific fluorescence images were obtained using confocal microscopy with a 405 nm excitation laser (Figure 2-7B). To test the labeling efficiency of the second-generation covalent TMP-tag, HEK 293T cells expressing eDHFR:L28C-6X His were incubated with media containing 1 μM A-TMP-Dapoxyl. At certain time points, cells were harvested, lysed, analyzed by SDS-PAGE and blotted with Anti-6X His. According to the band shift, near-quantitative labeling was achieved under the condition of 3 h incubation (Figure 2-7C). These labeling assays pave the ways toward the development of novel biophysical, physiological and multi-color pulse-chase applications with TMP-tag.

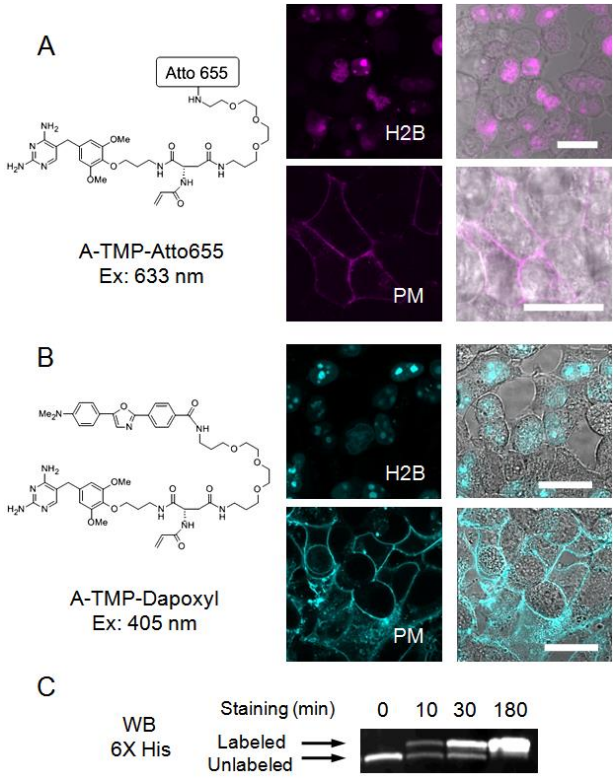



**Figure 2-7. Labeling of cellular protein targets with A-TMP-Atto655 and A-TMP-Dapoxyl.** (A) HEK293T cells transiently expressing H2B-eDHFR:L28C or PMLS-eDHFR:L28C were incubated

with 1 μM A-TMP-Atto 655 in media for 3 h, washed twice, and imaged using confocal and differential interference contrast (DIC) microscopy. Atto 655 was excited at 633 nm. (B) A-TMP-Dapoxyl was tested under the same condition as in (A), except Dapoxyl was excited at 405 nm. Scale bars, 25 μm. (C) Western blot analysis of the labeling efficiency. HEK293T cells transiently expressing eDHFR:L28C-6X His were incubated with 1 μM A-TMP-Dapoxyl in media for 10 min, 30 min and 3 h before lysed and analyzed by SDS-PAGE / Western blot with 6X His antibody.

## 2.4 Discussion

Together these results establish that by improving the design and hence the reaction half-life of our covalent TMP-tag, the v2.0 covalent TMP-tag is now a robust and general reagent for live cell imaging. The covalent TMP-tag design was improved by optimizing the spatial positioning of the Michael addition pair—the engineered Cys residue on eDHFR and the acrylamide conjugated to the A-TMP-probe heterodimer. While numerous chemical tags have been reported, our v2.0 covalent TMP-tag is one of the few examples that is selective enough to enable high signal to noise imaging of intracellular proteins, particularly diffuse, cytoplasmic proteins. Impressively, this speed and selectivity is achieved not from enzyme catalysis but rather from ligand-receptor binding followed by a proximity-induced organic reaction using a well-chosen mild electrophile. To date, the v2.0 covalent TMP-tag has been shown successful for a variety of protein targets and mammalian cell lines, and we expect it to be broadly useful to the community for imaging a wide range of proteins in living cells.

From a chemical perspective, proximity-induced reactions offer a combination of reactivity and specificity, which are both critical for protein labeling in live cells. Traditional protein-conjugation reagents such as maleimide electrophiles, which are designed for labeling purified proteins, cannot provide the desired selectivity for labeling proteins within the cell. More recently introduced bio-orthogonal reactions, such as the copper-less click reaction[37] or the photo-induced reactions which are triggered by UV-irradiation[38], require

unnatural amino acid incorporation[39], and are technically demanding and/or damaging to cells. In comparison, the non-covalent eDHFR-TMP interaction specifically accelerates the covalent reaction between the engineered Cys on eDHFR and the acrylamide electrophile. This approach, conceptually resembling the biomolecule-templated organic reactions[40], expands the scope of bioconjugation reactions as well as synthetic biology.

From an engineering point of view, proximity-induced reactions are facile implements for the development of novel chemical-tags. In the case of covalent TMP-tag, the specificity between ligand (A-TMP-Probe) and receptor (eDHFR) is guaranteed by high-affinity enzyme-inhibitor recognition as opposed to heavy-metal chelations (FlAsH/ReAsH)[5] or additional enzyme-catalyzed reactions (PRIME)[10]. Significantly, the covalent TMP-tag, which is based on high affinity binding, exhibits superior specificity and efficiency that enables intracellular protein labeling with minimal background. Using similar approaches, the vast pool of bioactive natural products and hit compounds from combinatorial libraries could be potentially engineered into orthogonal chemical tags based on proximity-induced reactions[41].

With an *in vitro* labeling half-life of 8 min, the second-generation covalent TMP-tag is significantly improved over our first-generation design[20]. Although the reaction rate is slower than the suicide-substrate based tags, e. g. SNAP tag[6], we consider the rate difference of little practical significance given that it typically requires 10 min to over an hour to label proteins in living cells, with uptake of the organic fluorophore considered to be the rate-limiting step. If needed, however, the labeling reaction kinetics could likely be further optimized by either molecular engineering of the small molecule ligand or directed evolution of eDHFR, or both. Notably, an advantage to a chemical tag based on high-affinity binding

is that it does not require the high concentration of ligand-probe conjugate necessary with enzyme-based chemical tags, where $K_M$s typically range from μM to mM, leading to high background noise from unbound fluorophore and necessitating extensive washing steps.

The second-generation covalent TMP-tag reported here is seen as a pressing improvement of the TMP-tag towards advanced protein-labeling applications. With its improved labeling kinetics and well-demonstrated cellular behavior, one might be able to track single protein molecules inside a cell[42] with a fluorophore of high photon-count. Moreover, the viability and robustness of the second-generation covalent TMP-tag point the way to multi-color protein labeling using orthogonal chemical tags.

## 2.5 Conclusion

By improving the reaction geometry of our covalent TMP-tag, we now have a v2.0 covalent TMP-tag that is a robust cellular reagent. This v2.0 covalent TMP-tag is an important addition to the limited arsenal of orthogonal covalent chemical tags available for multi-color imaging. Because our covalent TMP-tag is based on a modular organic reaction rather than a specific enzyme modification, we expect to be able to more readily build additional features into the covalent TMP-tag and generate new orthogonal tags simply by extending the Michael addition reaction to other drug-receptor pairs. While used here for live cell imaging, the covalent TMP-tag can be used broadly as a biotin-avidin surrogate for *in vitro* applications or in other applications of chemical dimerizers in live cells. Beyond the utility of proximity-induced reactivity for chemical tag engineering, the excellent reactivity and specificity of the proximity-induced Michael addition in a live cell illustrates the potential of organic chemistry for synthetic biology.

## 2.6 Supporting information and experimental methods

### 2.6.1 Molecular design

Using the model described in Figure 2-2, an approximation was made of the minimum linker length between TMP and the electrophile that would allow proximity-induced covalent labeling to occur upon binding to a mutant eDHFR containing each of these Cys mutants (Table 2-S1). The minimum length was approximated as the arc, $L = 2\pi R\sin\theta$, where $R$ is the radius of gyration and $\theta$ is the angle between the 4'-OH and the $\beta$-atom of each selected residue, using the center of the protein as the vertex.

**Table 2-S1: Estimation of the Minimum Linker Length for Proximity-Induced Covalent Labeling**

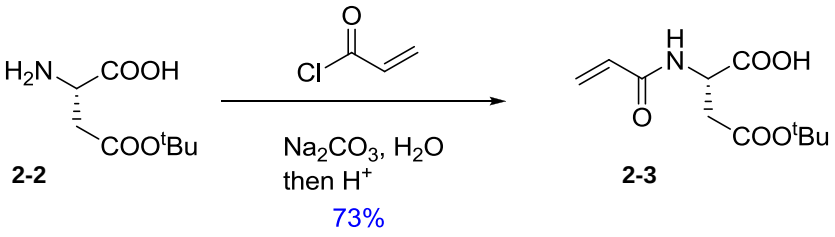
| Residue | Angle | Length (Å) |
|---|---|---|
| Ala19 | 38.72 | 10.27 |
| Asn23 | 30.83 | 8.18 |
| Leu28 | 35.62 | 9.45 |
| Arg52 | 35.86 | 9.51 |

### 2.6.2 Synthetic chemistry

**General Methods.** Unless otherwise noted, reagents and solvents were obtained from Aldrich and were used without further purification. Anhydrous N,N-dimethylformamide (DMF) was from Sure Seal$^{TM}$ bottles purchased from Aldrich.

Compound **2-5** is purchased form Ace Synthesis LLC. Nuclear magnetic resonance (NMR) spectra were recorded on a Bruker 500 (500MHz) Bruker 400 (400 MHz) or Bruker 300 (300 MHz) Fourier Transform (FT) NMR spectrometers at Columbia University, Chemistry Department. $^1$H NMR spectra are tabulated in the following order: multiplicity (s, singlet; d, doublet; t, triplet; m, multiplet; br, broad), number of protons. Fast Atom Bombardment (FAB) high resolution mass spectra (HRMS) were recorded on a JMS-HX110A mass spectrometer, and low resolution electron spray ionization (ESI) MS were recorded on a JMS-LC mate mass spectrometer.

Synthesis of carboxylic acid **2-3**:

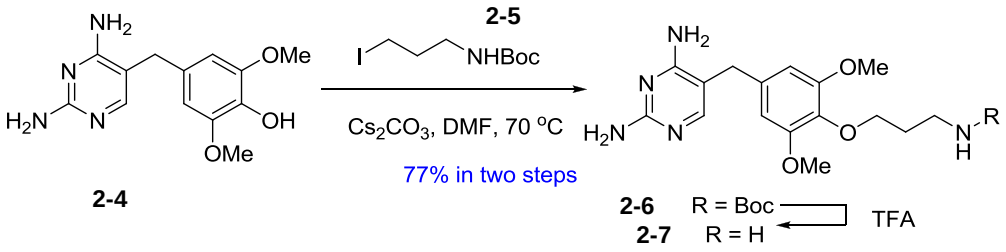



Acryloyl chloride (502 mg, 449 μL, 5.55 mmol) was added dropwise to a solution of Compound **2-2** (1.00 g, 5.28 mmol) and sodium carbonate (840 mg, 7.93 mmol) on an ice bath. The reaction was then warmed to RT and stirred for another 1 h. The pH of the solution was then adjusted to 2 by addition of 1M $NaHSO_4$ aqueous solution. Then the mixture was extracted by ethyl acetate (2 x 100 ml). The organic layers were combined, dried over anhydrous $Na_2SO_4$ and concentrated to yield compound 3 (0.93 g, 3.85 mmol, 73%) as a colorless oil.

$^1$H NMR (300 MHz, $CDCl_3$) δ ppm: 6.80 (d, J = 7.5 Hz, 1 H); 6.35 (dd, J = 16.8 Hz, 1.2 Hz, 1 H); 6.16 (dd, J = 16.8 Hz, 10.2 Hz, 1 H); 5.75 (dd, J = 10.2 Hz, 1.2 Hz, 1 H); 4.89(m, 1 H); 3.02 (dd, J = 17.1 Hz, 4.2 Hz, 1 H); 2.78 (dd, J = 17.1 Hz, 5.4 Hz, 1 H); 1.44 (s, 9 H)

HRMS (FAB+) m/z Calcd. for C11H18O5N [M+H]+:244.1185. Found: 244.1181

Synthesis of TMP-amine **2-7**:

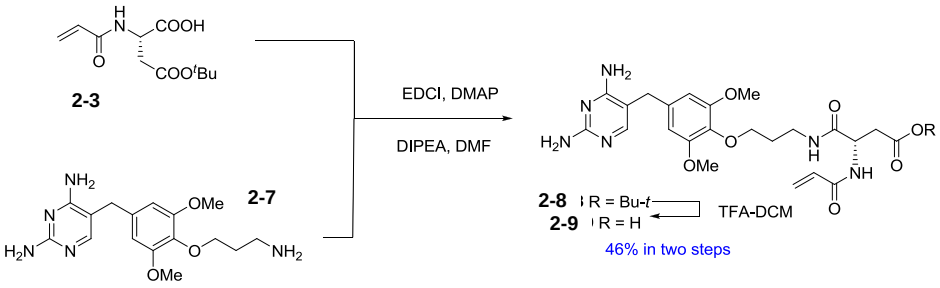



Compound **2-4** (600 mg, 2.16 mmol), compound **2-5** (618 mg, 2.16 mmol) and cesium carbonate (1.41 g, 4.32 mmol) were dissolved in anhydrous DMF (30 ml). The reaction was allowed to stand at 70 ℃ for 7 h, followed by the removal of solvent *in vacuo*. The crude mixture was purified by column chromatography with silica gel (elutant: 10% MeOH in DCM) to yield compound **2-6**, which was deprotected by directly dissolving in 10 ml TFA while stirring. After 4 h, the TFA was removed *in vacuo* to yield compound **2-7** in the form of its TFA salt (715 mg, 1.66 mmol, 77%) as colorless oil.

$^1$H NMR (300 MHz, $CD_3OD$) δ ppm: 7.28 (s, 1 H); 6.64 (s, 2 H); 4.12 (t, *J* = 5.4 Hz, 2 H); 3.87 (s, 6 H); 3.70 (s, 2 H); 3.30 (t, *J* = 6.0 Hz, 2 H); 2.08 (m, 2 H)

HRMS (FAB+) m/z Calcd. for $C_{16}H_{24}O_3N_5$ $[M+H]^+$: 334.1879. Found: 334.1883

Synthesis of A-TMP-COOH **2-9**:



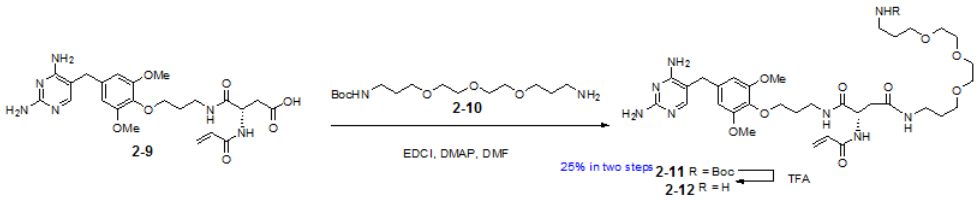

N,N-Diisopropylethylamine (259 mg, 348 μl, 2.00 mmol) was added dropwise to a solution of compound **2-3** (187 mg, 0.77 mmol), compound **7** (271 mg, 0.63 mmol, in the

form of TFA salt), 4-Dimethylaminopyridine (39 mg, 0.30 mmol) and 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (241 mg, 1.26 mmol) in 10 ml anhydrous DMF at RT. The reaction was stirred for 20 h at RT and concentrated. The mixture was then purified by silica gel flash chromatography (10% MeOH in DCM) to give the coupling product **2-8**, which was directly dissolved in 20 ml TFA-DCM (1:1, v/v) and stirred for 4 h at RT. The solvent was removed *in vacuo* and the crude product was purified by column chromatography on silica gel (30% MeOH in DCM) to give carboxylic acid **2-9** (127 mg, 253 μmol, 46%) as a colorless oil.

$^1$H NMR (300 MHz, $CD_3OD$) δ ppm: 7.25 (s, 1 H); 6.57 (s, 2 H); 6.24 (dd, *J* = 17.1 Hz, 9.0 Hz, 1 H); 6.16 (dd, *J* = 17.1 Hz, 3.0 Hz, 1 H); 5.61 (dd, *J* = 9.0 Hz, 3.0 Hz, 1 H); 4.80 (m, 1 H); 3.98 (m, 2 H); 3.82 (s, 6 H); 3.68 (s, 2 H); 3.46 (t, *J* = 6.0 Hz, 2 H); 2.87 (dd, *J* = 16.8 Hz, 6.0 Hz, 1 H); 2.72 (dd, *J* = 16.8 Hz, 7.5 Hz, 1 H); 1.88 (m, 2 H)

HRMS (FAB+) m/z Calcd. for $C_{23}H_{31}O_7N_6$ $[M+H]^+$:503.2254. Found: 503.2273

Synthesis of A-TMP-PEG-$NH_2$, compound **2-12**:



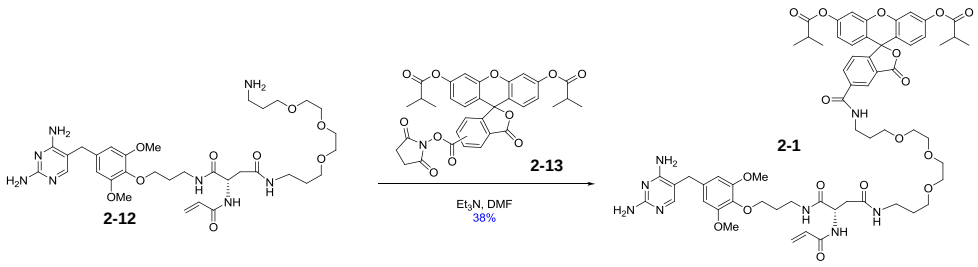

To the solution of compound **2-9** (40 mg, 80 μmol), compound **2-10** (38 mg, 119 μmol), 4-dimethylaminopyridine (5 mg, 40 μmol) in 10 ml anhydrous DMF was added 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (31 mg, 160 μmol) at RT. The reaction was stirred for 20 h at RT before concentrated. The mixture was then purified by silica gel flash chromatography (10% MeOH in DCM) to give the coupling product **2-11**, which was directly dissolved in 5 ml trifluoroacetic acid (TFA) and stirred for 4 h at RT. The solvent

was removed *in vacuo* and the amine product was further purified by column chromatography on reverse phase C18 silica gel with 10% ACN in water to give A-TMP-PEG-$NH_2$, compound **2-12** (16 mg, 20 μmol, in the form of TFA salt, 25%) as white solid.

$^1$H NMR (400 MHz, $CD_3OD$) δ ppm: 7.25 (s, 1 H); 6.57 (s, 2 H); 6.22 (dd, *J* = 17.2 Hz, 9.6 Hz, 1 H); 6.15 (dd, *J* = 17.2 Hz, 2.8 Hz, 1 H); 5.61 (dd, *J* = 9.6 Hz, 2.8 Hz, 1 H); 4.83 (m, 1 H); 3.96 (m, 2 H); 3.81 (s, 6 H); 3.69-3.62 (m, 10 H); 3.59-3.55(m, 2 H); 3.49 (t, *J* = 6.0 Hz, 2 H); 3.48-3.40 (m, 2 H); 3.24 (dt, *J* = 7.2 Hz, 3.2 Hz, 2 H); 3.10 (t, *J* = 6.4 Hz, 2 H); 2.76 (dd, *J* = 14.8 Hz, 5.6 Hz, 1 H); 2.58 (dd, *J* = 14.8 Hz, 7.6 Hz, 1 H); 1.92 (m, 2 H); 1.87 (m, 2 H); 1.73 (m, 2 H)

HRMS (FAB+) m/z Calcd. for $C_{33}H_{53}O_9N_8$ $[M+H]^+$:705.3936. Found: 705.3917

Synthesis of A-TMP-PEG-Fl, compound **2-1**:




5 μL triethylamine was added to a solution of compound **2-12** (3.0 mg, 3.7 μmol) and compound **2-13** (4.0 mg, 6.5 μmol, mixture of isomers) in 1 ml anhydrous DMF. The mixture was stirred at RT for 12 h before concentrated. Reaction residue was purified by reverse phase HPLC to give compound **2-1** (1.7 mg, 1.4 μmol, 38%) as a single isomer.

HPLC condition: starting with 33%: 67% acetonitrile: Water, gradient elution for 50 min, end with 50%:50% ACN:Water. Retention time: 42-43 min.

$^1$H NMR (400 MHz, $CD_3OD$) δ ppm: 8.46 (d, *J* = 0.8 Hz, 1 H ), 8.22 (dd, *J* = 8.0 Hz, 1.6 Hz, 1 H); 7.38(d, *J* = 8.0 Hz, 1 H); 7.23 (s, 1 H); 7.19 (m, 2 H); 6.90 (m, 4 H); 6.22 (dd,

*J* = 16.8 Hz, 9.6 Hz, 1 H); 6.14(dd, *J* = 16.8 Hz, 2.4 Hz, 1 H); 5.59 (dd, *J* = 9.6 Hz, 2.4 Hz, 1 H); 4.78 (t, *J* = 6.4 Hz, 1 H); 3.94 (t, *J* = 5.6 Hz, 2 H); 3.79 (s, 6 H); 3.66-3.60 (m, 10 H); 3.55-3.52 (m, 4 H); 3.47-3.40 (m, 4 H); 3.20 (m, 2 H); 2.85 (m, 2 H); 2.71 (dd, *J* = 15.2 Hz, 6.0 Hz, 1 H); 2.60 (dd, *J* = 15.2 Hz, 7.2 Hz, 1 H); 1.92 (m, 2 H); 1.83 (m, 2 H); 1.68 (m, 2 H); 1.30 (d, *J* = 7.2 Hz, 12 H)

MS (FAB+) m/z Calcd. for $C_{62}H_{75}O_{17}N_8$ $[M+H]^+$:1203.53 Found:1203.63

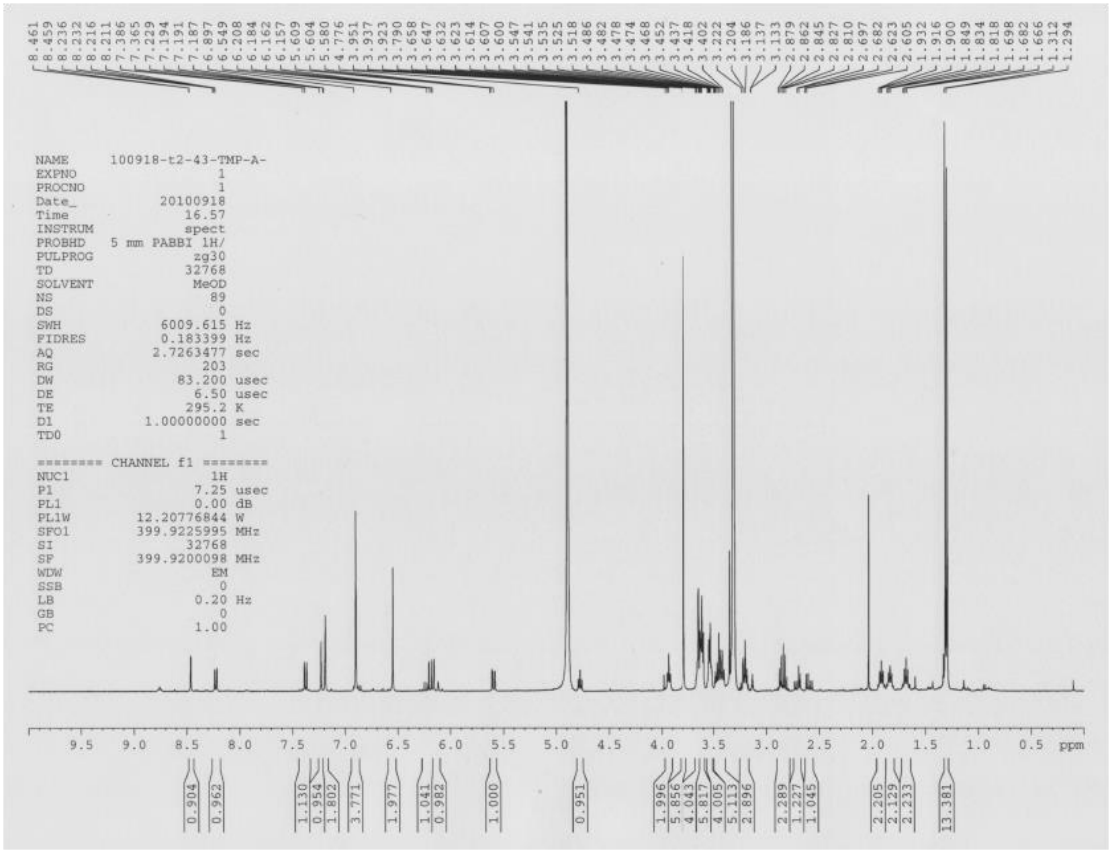
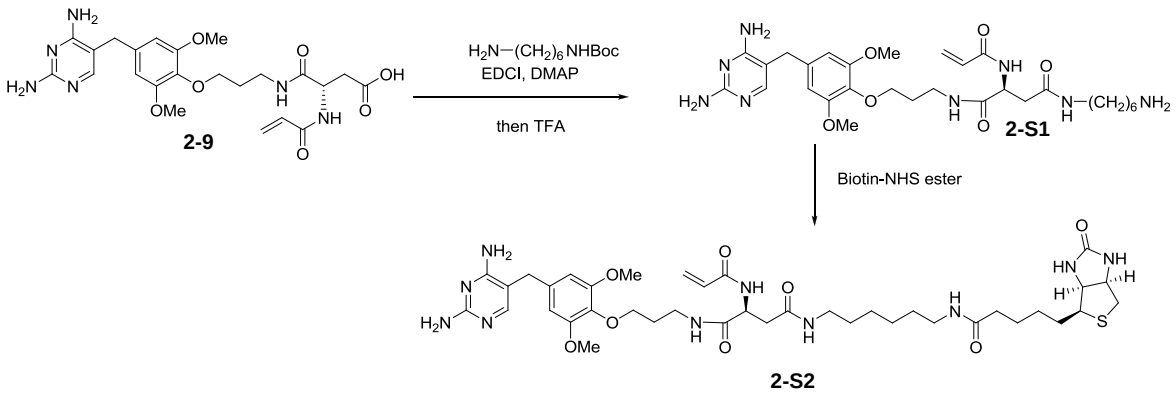



**Figure 2-S1. $^1$H NMR spectrum of compound 2-1.**

A-TMP-Biotin heterotrimer, Compound **2-S2**, is prepared following the synthetic plan shown in Scheme 2-S1.




**Scheme 2-S1. Synthetic route toward Compound 2-S2, A-TMP-Biotin heterotrimer**

Synthesis of compound **2-S1**:

To the solution of compound **2-9** (82 mg, 163 μmol), *N*-Boc-1,6-hexanediamine hydrochloride (62 mg, 250 μmol), 4-dimethylaminopyridine (10 mg, 80 μmol) in 5 ml anhydrous DMF was added 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (62 mg, 326 μmol) and *N*,*N*-Diisopropylethylamine (63 mg / 85 μL, 490 μmol) at RT. The reaction was stirred for 20 h at RT before concentrated *in vacuo*. The mixture was then purified by silica gel flash chromatography (10% MeOH in DCM) to give the coupling product. The coupling product was directly dissolved in 5 ml trifluoroacetic acid (TFA) and stirred for 4 h at RT. The solvent was removed *in vacuo* and a fraction of the amine product was further purified by reverse phase HPLC to give A-TMP-C6-$NH_2$, compound **2-S1** (12.0 mg, in the form of TFA salt) as white solid.

HPLC condition: starting with 10%: 90% acetonitrile: Water, gradient elution for 50 min, end with 40%:60% ACN:Water. Retention time: 13-15 min.

$^1$H NMR (300 MHz, $CD_3OD$) δ ppm: 7.25 (s, 1 H); 6.57 (s, 2 H); 6.22 (dd, *J* = 17.1 Hz, 9 Hz, 1 H); 6.14 (dd, *J* = 17.1 Hz, 3 Hz, 1 H); 5.61 (dd, *J* = 9 Hz, 3 Hz, 1 H); 4.81 (m, 1 H); 3.96 (m, 2 H); 3.81 (s, 6 H); 3.67 (s, 2 H); 3.48-3.4(m, 2 H); 3.16 (t, *J* = 6.9 Hz, 2 H);

2.91 (t, *J* = 7.5 Hz, 2 H); 2.74 (dd, *J* = 15 Hz, 6 Hz, 1 H); 2.58 (dd, *J* = 15 Hz, 7.5 Hz, 1 H); 1.87 (m, 2 H); 1.64 (m, 2 H); 1.5 (m, 2 H); 1.37 (m, 4 H)

HRMS (FAB+) m/z Calcd. for $C_{29}H_{45}O_6N_8$ $[M+H]^+$:601.3457. Found: 601.3459

Synthesis of compound **2-S2**, A-TMP-Biotin:

5 μL *N*,*N*-Diisopropylethylamine was added to a solution of compound **2-S1** (5.4 mg, 9 μmol) and Biotin-NHS ester (6 mg, 18 μmol) in 1 ml anhydrous DMF. The mixture was stirred at RT for 12 h before concentrated. Reaction residue was purified by preparative TLC to give compound **2-S2** A-TMP-Biotin.

$^1$H NMR (400 MHz, $CD_3OD$) δ ppm: 7.52 (s, 1 H); 6.52 (s, 2 H); 6.19 (dd, *J* = 17.2 Hz, 8.6 Hz, 1 H); 6.14 (dd, *J* = 17.2 Hz, 2.4 Hz, 1 H); 5.59 (dd, *J* = 8.6 Hz, 2.4 Hz, 1 H); 4.79 (m, 1 H); 4.48 (dd, *J* = 8 Hz, 4.4 Hz, 1 H); 4.29 (dd, *J* = 8 Hz, 4.4 Hz, 1 H); 3.95 (m, 2 H); 3.79 (s, 6 H); 3.64 (s, 2 H); 3.48-3.4(m, 2 H); 3.19 (m, 1 H); 3.17-3.1 (m, 4 H); 2.92 (dd, *J* = 12.8 Hz, 4.8 Hz, 1 H); 2.72 (dd, *J* = 15.2 Hz, 6 Hz, 1 H); 2.70 (dd, *J* = 12.8 Hz, 4.4 Hz, 1 H); 2.59 (dd, *J* = 15.2 Hz, 7.6 Hz, 1 H); 2.18 (t, *J* = 7.2 Hz, 2 H); 1.86 (m, 2 H); 1.78-1.55 (m, 4 H); 1.52-1.38 (m, 6 H); 1.37-1.26 (m, 6 H)

HRMS (FAB+) m/z Calcd. for $C_{39}H_{59}O_8N_{10}S$ $[M+H]^+$:827.4238. Found:827.4276

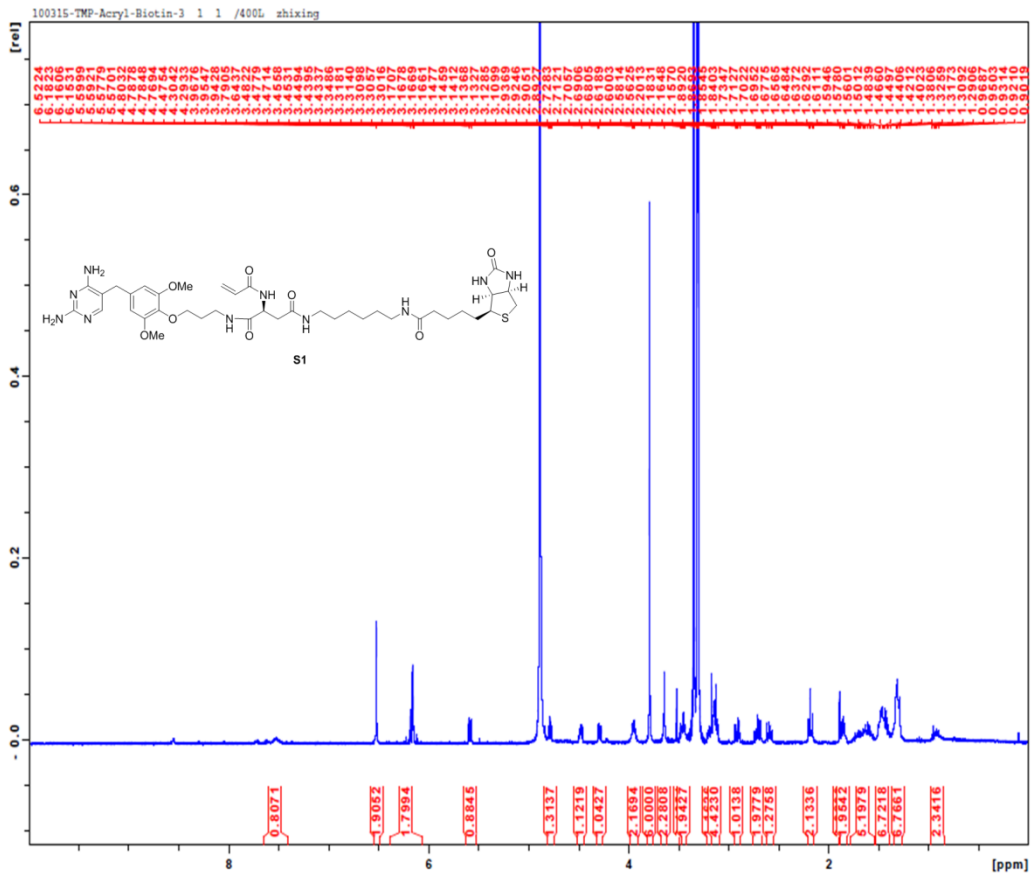



**Figure 2-S2. $^{1}$H NMR spectrum of compound 2-S1.**

Synthesis of A-TMP-Dapoxyl

5 μL triethylamine was added to a solution of compound **2-12** (2 mg, 2.4 μmol) and Dapoxyl-NHS (1 mg, Invitrogen D-10161) in 1 ml anhydrous DMF. The mixture was stirred at RT overnight before concentrated *in vacuo*. Reaction residue was purified by reverse phase HPLC to give A-TMP-Dapoxyl (1.3 mg, 1.3 μmol, ~50 %)

$^{1}$H NMR (400 MHz, $CD_3OD$) δ ppm: 8.14 (d, *J* = 8.4 Hz, 2H); 7.96 (d, *J* = 8.4 Hz, 2H); 7.70 (d, *J* = 8.0 Hz, 2H); 7.45 (s, 1 H); 7.22 (s, 1 H); 6.95 (d, *J* = 7.2 Hz, 2H); 6.54 (s, 2 H); 6.22 (dd, *J* = 17.2 Hz, 9.6 Hz, 1 H); 6.14 (dd, *J* = 17.2 Hz, 2.4 Hz, 1 H); 5.60 (dd, *J* = 9.6 Hz, 2.8 Hz, 1 H); 4.77 (m, 1 H); 3.95 (m, 2 H); 3.80 (s, 6 H); 3.69-3.59 (m, 10 H); 3.57-

3.41(m, 6 H); 3.21 (dt, *J* = 7.6 Hz, 1.6 Hz, 2 H); 3.06 (s, 6 H); 2.70 (dd, *J* = 15.2 Hz, 6.4 Hz, 1 H); 2.59 (dd, *J* = 15.2 Hz, 7.2 Hz, 1 H); 1.90 (m, 2 H); 1.85 (m, 2 H); 1.69 (m, 2 H)

HRMS (FAB+) m/z Calcd. for $C_{51}H_{67}O_{11}N_{10}$ $[M+H]^+$:995.4985. Found: 995.5025

Synthesis of A-TMP-Atto655

1 μL triethylamine was added to a solution of compound **2-12** (1 mg, 1 μmol) and Atto655-NHS (1 mg, Sigma 76245) in 1 ml anhydrous DMF. The mixture was stirred at RT overnight before concentrated *in vacuo*. Reaction residue was purified by reverse phase HPLC to give A-TMP-Atto655 (1.2 mg, 1 μmol).

MS (FAB+) m/z Calcd. for $C_{60}H_{83}O_{14}N_{11}S$ $[M+H]^+$:1214.59 Found:1214.82

### 2.6.3 Vector construction, protein expression and purification

Construction of eDHFR variant vectors for protein purification

The vector encoding eDHFR:L28C for *E.coli* over-expression and protein purification was previously published p2250. The intermediate vector p2247 encoding eDHFR-2C was also used to construct eDHFR:A19C, eDHFR:N23C, and eDHFR:R52C by means of site-directed mutagenesis (Strategene'sQuikChange Mutagenesis Kit). Primers for mutagenesis: 5'-GTT ATC GGC ATGGAA AAC TGC ATG CCG TGG AAC CTG CC-3' and 5'- GGC AGG TTC CAC GGC ATG CAG TTT TCC ATG CCG ATA AC-3' (C85S, C142S, and A19C, p2248); 5'-CGCCAT GCC GTG GTG CCT GCC TGC CGA T -3' and 5'- ATC GGC AGG CAG GCA CCA CGG CAT GGC G-3' (C85S, C142S, and N23C, p2249); 5'-CTGGGA ATC AAT CGG GTG CCC GTT GCC AGG ACG C-3' and 5'- GCG TCC TGG CAA CGG GCA CCC GAT TGA TTC CCA G-3' (C85S, C142S, and R52C, p2250). The resulting coding regions were sequenced in full.

Protein Purification

The resulting eDHFR:L28C, eDHFR:A19C, eDHFR:N23C, and eDHFR:R52C plasmids were expressed in BL21(DE3) pLysS cells (Invitrogen). Cells were grown at 37 ℃ to an $OD_{600}$ of 0.6, induced with 0.4 mM IPTG for three hours and purified using nickel-NTA spin columns (Qiagen). The proteins were dialyzed three times in phosphate buffered saline (PBS) at 4 ℃, snap frozen and stored at −80°C.

Construction of H2B-eDHFR:L28C

The gene encoding eDHFR was amplified using PCR from the previously published p1008 vector with primers 5'-ACGTCACCGGTCGCCACCATGGTGGGTTCTGGTGGTTCTGGTATCAGTCTGATT GCGGCG-3' (AgeI, coding strand) and 5'-ACGTCGCGGCCGCTTTAGTGATGGTGATGGTGATGCCGCCGCTCCAGAATCT -3' (NotI, noncoding strand). The fragment was then inserted between the AgeI and NotI sites of vector encoding H2B-EGFP (Addgene p11680) replacing the EGFP gene. The resulting vector (p2606) encoding H2B-eDHFR was then subject to site-directed mutagenesis (SDM) to give H2B-eDHFR:L28C (p2607). Primers for SDM: 5'- GGA ACC TGC CTG CCG ATT GCG CAT GGT TTA AAC GCA AC -3'; 5'- GTT GCG TTT AAA CCA TGC GCA ATC GGC AGG CAG GTT GC-3'. The resulting coding region was sequenced in full.

Construction of TOMM20-eDHFR:L28C

The parent vector pSNAP-Cox8A was purchased from New England Biolab. The gene of eDHFR was amplified from p1008 using primers 5' -GCA TAC GTC GAT ATC AAG CTT ACC ATG ATC AGT CTG ATT GCG G -3' (EcoRV, coding strand) and 5'-GCA TAC GTC CTC GAG TTA CCG CCG CTC CAG AA - 3' (XhoI, noncoding strand),

and then inserted between the EcoRV and XhoI sites of pSNAP-Cox8A, creating p2492. Human Tomm20 gene (Invitrogen) was amplified with primers 5’ -GTC GAC ATC GAT ATG GTG GGT CGG AAC AG -3’ (ClaI, coding strand) and 5’- GCA GTC GAT ATC ATC TAC GAC ATC TT - 3’ (EcoRV, noncoding strand), and then inserted between the ClaI and EcoRV sites of p2492 to give p2605 encoding TOMM20-eDHFR. Subsequent SDM on p2605 yielded TOMM20-eDHFR:L28C (p2608). The mutagenesis primers were the same as for H2B-eDHFR:L28C.

Construction of MLC-eDHFR:L28C

The parent vector encoding MLC-eDHFR was previously published. The eDHFR gene in this vector was changed to eDHFR:L28C by SDM using the same set of mutagenesis primers as for eDHFR:L28C expression vector.

Construction of $\alpha$-Actinin-eDHFR:L28C

Vector p2606 encoding H2B-eDHFR was used as the parent vector. Human α-Actinin gene (Addgene) was amplified with primers 5’- CGA ATT CTG CAG TCG ACG GTA CCG CCA TGG ACC ATT ATG ATT CTC AGC AAA CC -3’ (SalI, coding strand) and 5’- CAT GGT GGC GAC CGG TGG ATC GAG GTC ACT CTC GCC GTA CA-3’ (AgeI, noncoding strand), and then inserted between the SalI and AgeI sites of p2606. The resulting plasmid (p2611) was subject to SDM to create $\alpha$-Actinin-eDHFR:L28C vector (p2612). The mutagenesis primers were the same as for H2B-eDHFR:L28C.

Construction of eDHFR:L28C-6xHis

Vector p2607 encoding H2B-eDHFR:L28C was used as the parent vector. eDHFR:L28C gene was amplified with primers 5’-taagacggtaccATGATCAGTCTGATTGCGGCG -3’ (KpnI, coding strand)

and 5'-ACGTCGCGGCCGCTTTAGTGATGGTGATGGTGATGCCGCCGCTCCAGAATCT -3' (NotI, noncoding strand), and then inserted between the KpnI and NotI sites of p2607. The resulting coding region was sequenced in full.

### 2.6.4 *In vitro* alkylation

Purified eDHFR mutants (5 μM) was incubated with A-TMP-B (10 μM) in PBS with reduced glutathione (1 mM) at 37 ℃, with or without NADPH (50 μM). At selected time points, aliquots (20 μL) were removed from the reaction mixture, quenched with 6X SDS and boiled for 5 minutes. Samples from the in vitro alkylation experiments were analyzed by SDS-PAGE on Criterion 15% Tris-HCl gels (BioRad) for 60 min at 200 V. Bands were quantified by densitometry analysis of Coomassie stained gels by Image-J.

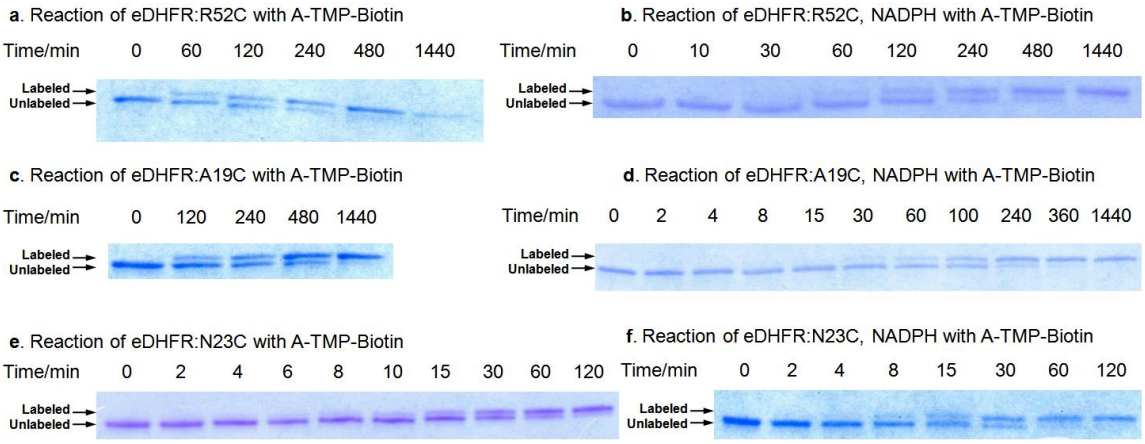



**Figure 2-S3. SDS-PAGE analysis of *in vitro* alkylation aliquots from different substrates.**

### 2.6.5 Cell culture, transfection and labeling

HEK 293T cells were cultured in DMEM w/ glutamine (Gibco #11995) with 10% v/v fetal bovine serum and 1% v/v Pen/Strep. All cells were maintained under 5% $CO_2$ at

37 ℃. For live cell protein labeling, cells were plated in 8-well chambered #1 borosilicate coverglass (Thermo, Nunc 155411) 24 h before transfected with expression plasmids for eDHFR-L28C fused protein of interest(s) (0.4 μg DNA for one well) using Fugene HD (Roche).

MEF cells were cultured in DMEM (Gibco #11965) with 10% v/v fetal bovine serum, 1.5% v/v HEPES and 1% v/v Pen/Strep. All cells were maintained under 5% $CO_2$ at 37 ℃. Transient transfection of plasmids was performed 2 days before labeling experiment through electroporation using a Nucleofector (Lonza) according to the manufacturer's protocol (MEF-1solution and program T-20).

24 h after transfection, 300 μL media with 0.3 μL A-TMP-Fl stock solution (1 mM in DMF), a final concentration of 1 μM, was added to the well. Cells were incubated with this staining solution for 10 min at 37 ℃, followed by washing with fresh media for 2 times before imaging or cellular analysis.

### 2.6.6 Imaging

Confocal images were obtained using LEICA TCS SP5 confocal microscope with a HCX PL APO CS 20.0x0.70 DRY UV objective or a HCX PL APO CS 100.0x1.46 OIL objective. Images were processed by LAS AF software. Confocal images under 405 nm excitation were obtained using Zeiss LSM 710 with a PL APO 63x1.40 Oil objective. Images were processed by ZEN software.

### 2.6.7 In-gel fluorescence/Western blotting analysis.

After the 10-min treatment of labeling reagent, cells were harvested with trypsin/EDTA. Trypsinization was stopped by addition of DMEM and the cell density was

determined to be around $5x10^5$ cells per wall. The cells were then pelleted by centrifugation at 1000 rpm for 4 minutes. The supernatant was removed, and the pellet was washed with 1 mL of PBS per $5x10^5$ cells and repelleted. Again, the supernatant was removed and the pellet was resuspended in lysis buffer to a concentration of $5x10^5$ cells in 12.5 µL of lysis buffer and lysed on ice for 20 minutes. The lysate was then centrifuged at 13000 rpm for 10 minutes and the supernatant was mixed with loading buffer and heated to 95° C for 5 minutes. Samples from covalent labeling experiments were run on Criterion 15% Tris-HCl gels (BioRad) for 60 min at 200 V. The proteins were detected by in-gel fluorescence scanning using a Typhoon™ Trio scanner. The gel was analyzed for green fluorescence by scanning with the 488 nm lasers and an emission filter of 520 (BP 40). The PMT value was adjusted and a high-resolution scan was collected at 100 µm resolution.

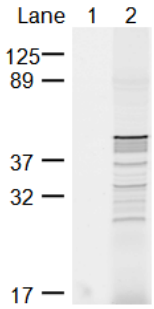



**Figure 2-S4. In-gel fluorescence analysis.** Labeled HEK 293T cells were lysed and then analyzed by SDS-PAGE and in-gel fluorescence scanning with an excitation laser at 488 nm: Lane 1, non-transfected HEK 293T cells; Lane 2, HEK293T cells transfected with a vector encoding the H2B-eDHFR:L28C fusion protein.

Western blotting experiments were performed under Odyssey western blotting protocol (Li-Cor) using an iBlot® Gel Transfer Device (Invitrogen) and the corresponding PVDF membrane. Primary/secondary antibodies: Anti-H2B (cell signaling #2722, 1: 1000) / IRDye 800CW Goat anti-Rabbit IgG (H + L) (LI-COR 926-32211) ; Anti-6XHis

(Invitrogen R932-25, 1: 5000) / Alexa Fluor® 680 Goat Anti-Mouse IgG (H+L) (Invitrogen A-21058). The final western blot membranes were scanned with Odyssey scanner (LI-COR).

## 2.7 References


1. Jing, C.; Cornish, V. W. *Acc. Chem. Res.*, **2011**, *44,* 784
2. Hinner, M. J.; Johnsson, K. *Curr. Opin. Biotechnol.* **2010**, *21*, 766.
3. Fernández-Suárez, M.; Ting, A. Y. *Nat. Rev. Mol. Cell. Biol.* **2008**, *9*, 929.
4. Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2006**, *312*, 217.
5. Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269.
6. Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat. Biotechnol.* **2003**, *21*, 86.
7. Gautier, A.; Juillerat, A.; Heinis, C.; Correa, I. R. Jr.; Kindermann, M.; Beaufils, F.; Johnsson, K. *Chem. Biol.* **2008**, *15*, 128.
8. Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; and Wood, K. V. *ACS Chem. Biol.* **2008**, *3*, 373.
9. Miller, L. W.; Cai, Y.; Sheetz, M. P.; and Cornish, V. W. *Nat. Methods* **2005**, *2*, 255.
10. Uttamapinant, C.; White, K. A.; Baruah, H.; Thompson, S.; Fernández-Suárez, M.; Puthenveetil, S.; and Ting, A. Y. *Proc. Natl. Acad. Sci.U.S.A.* **2010**, 107, 10914.
11. Hoskins, A. A.; Friedman, L. J.; Gallagher, S. S.; Crawford, D. J.; Anderson, E. G.; Wombacher, R.; Ramirez, N.; Cornish, V. W.; Gelles, J.; Moore, M. J. *Science* **2011**, *331*, 1289.
12. Long, M. J. C.; Pan, Y.; Lin, H.-C.; Hedstrom, L.; Xu, B. *J. Am. Chem. Soc.* **2011**, *133,* 10006
13. Zou, P.; Ting, A. Y. *ACS Chem. Biol.* **2011**, *6*, 308.
14. Wombacher, R.; Heidbreder, M.; van de Linde, S.; Sheetz, M. P.; Heilemann, M.; Cornish, V. W.; Sauer, M. *Nat. Methods* **2010**, *7*, 717.
15. Baccanari, D. P.; Daluge, S.; King, R. W. *Biochemistry* **1982**, *21*, 5068
16. Mattews, D. A.; Bolin, J. T.; Burridge, J. M.; Filman, D. J.; Volz, K. W.; Kaufman, B. T.; Beddell, C. R., Champness, J. N., Stammers, D. K.; Kraut, J. *J. Biol. Chem.* **1985**, *260*, 381
17. Calloway, N. T.; Choob, M.; Sanz, A.; Sheetz, M. P.; Miller, L. W.; Cornish, V. W. *ChemBioChem* **2007**, *8*, 767.

18. Gallagher, S. S.; Jing C.; Peterka, D. S.; Konate, M.; Wombacher, R.; Kaufman, L. J.; Yuste, R.; Cornish, V. W. *ChemBioChem* **2010**, *11*, 782.

19. Cai, Y.; Rossier, O.; Gauthier, N. C.; Biais, N.; Fardin, M.-A.; Zhang, X.; Miller, L. W.; Ladoux, B.; Cornish, V. W.; Sheetz, M. P. *J. Cell Sci.*, **2010**, *123*, 413-423

20. Gallagher, S. S.; Sable, J. E.; Sheetz, M. P.; Cornish V. W. *ACS Chem. Bio.* **2009**, *4*, 547

21. Breslow, R.; Corcoran, R.; Dale, J. A.; Liu, S.; Kalicky, P. *J. Am. Chem. Soc.*, **1974**, *96*, 1973

22. Breslow, R.; Corcoran, R. J.; Snider, B. B.; Doll, R. J.; Khanna, P. L.; Kaleya, R. *J. Am. Chem. Soc.*, **1977**, *99* , 905

23. Powers, J. C.; Asgian, J. L.; Ekici, Ö. D.; James, K. E. *Chem. Rev.*, **2002**, *102* , 4639

24. Fry, D. W.; Bridges, A. J.; Denny, W. A.; Doherty, A.; Greis, K. D.; Hicks, J. L.; Hook, K. E.; Keller, P. R.; Leopold, W. R.; Loo, J. A.; McNamara, D. J.; Nelson, J. M.; Sherwood, V.; Smaill, J. B.; Trumpp-Kallmeyer, S.; Dobrusin, E. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12022

25. Evans, M. J.; Cravatt, B. F. *Chem. Rev.* **2006**, *106*, 3279.

26. Levitsky, K.; Boersma, M. D.; Ciolli, C. J.; Belshaw, P. J. *ChemBioChem* **2005**, *6*, 890.

27. Krusemark, C. J.; Belshaw, P. J. *Org. Biomol. Chem.* **2007**, *5*, 2201.

28. Sawaya, M. R.; Kraut, J. *Biochemistry* **1997**, *36*, 586.

29. Polshakov, V. I.; Smirnov, E. G.; Birdsall, B.; Kelly, G.; Feeney, J. *J. Biomol. NMR* **2002**, *24*, 67.

30. *Maestro* version 9.1. Schrödinger, LLC, New York, NY.

31. DeLano, W. L. *The PyMOL Molecular Graphics System, DeLano Scientific*, Palo Alto, CA, **2002**

32. Izumi, S.; Urano, Y.; Hanaoka, K.; Terai, T.; Nagano, T. *J. Am. Chem. Soc.*, **2009**, *131*, 10189

33. Veech, R. L.; Eggleston, L. V. Krebs, H. A. *Biochem. J.* **1969**, *115*, 609.

34. Reiss, P. D.; Zuurendonk, P. F.; Veech, R. L. *Anal. Biochem.* **1984**, *140*, 162.

35. Diwu, Z.; Zhang, C.; Klaubert, D. H.; Haugland, R. P. *J. Photochem. Photobiol. A* **2000**, *131*, 95

36. Pham, W,; Pantazopoulos, P.; Moore, A. *J. Am. Chem. Soc.*, **2006**, *128*, 11736

37. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.*, **2004**, *126*, 15046

38. Lim, R. K. V.; Lin, Q. *Acc. Chem. Res.*, **2011**, *44,* 828

39. Hao, Z.; Hong, S.; Chen, X.; Chen, P. R. *Acc. Chem. Res.*, **2011**, *44,* 742

40. Li, X.; Liu, D. R. *Angew. Chem. Int. Ed.* **2004**, *43,* 4848

41. Choi, S.; Ong, D. S. T.; Kelly, J. W. *J. Am. Chem. Soc.*, **2010**, *132*, 16043

42. Li, G. -W.; Xie, X. S. *Nature*, **2011**, *475,* 308

# Chapter 3

# Mapping protein-specific micro-environments in live cells by fluorescence lifetime imaging of a hybrid genetic-chemical molecular rotor tag

The content of this chapter has been published in:

Gatzogiannis, E.; Chen, Z.; Wei, L.; Wombacher, R.; Kao, Y.-T.; Yefremov, G.; Cornish, V. W.; Min, W., “Mapping protein-specific micro-environments in live cells by fluorescence lifetime imaging of a hybrid genetic-chemical molecular rotor tag”. *Chem. Commun. (Camb).* **2012**, *48*, 8694. E.G. and Z.C. contributed equally to this work.

## 3.1 Chapter outlook

The micro-viscosity and molecular crowding experienced by specific proteins can regulate their dynamics and function within live cells. Taking advantage of the emerging TMP-tag technology, we present the design, synthesis and application of a hybrid genetic-chemical molecular rotor probe whose fluorescence lifetime can report protein-specific micro-environments in live cells.

Evangelos Gatzogiannis, I and Wei Min are the main contributors of this work. Wei Min conceived the concept. Evangelos Gatzogiannis aligned the fluorescence lifetime microscope with the help of Ya-Ting Kao. I synthesized the TMP-Cy3 probe based on the unpublished results of Richard Wombacher, Gregory Yefremov and Virginia Cornish. I and Lu Wei performed the cell culture experiments. Evangelos Gatzogiannis and I performed the *in vitro* measurements and cell imaging experiments. Evangelos Gatzogiannis, I and Wei Min wrote the paper with contributions from all co-authors.

## 3.2 Introduction

The micro-mechanical environment of proteins inside live cells and its effect on key biochemical processes are important yet unresolved issues in cell biology.[1] Intracellular viscosity plays an important role in biochemical processes such as signal transduction, nuclear envelope function, chromatin localization, ribonucleoprotein assembly and diffusion of reactive oxygen species.[2–4] Changes in viscosity at a sub-cellular level have been related to a number of diseases and pathologies.[5] The determination of local viscosities and other micro-environmental parameters within the nucleus and other critical cellular organelles in live cells is of great interest.[2–13]

Experimentally, intracellular viscosity has been measured by tracking fluorophore diffusion, fluorescence correlation spectroscopy, and fluorescence recovery after photobleaching.[6–9] These methods, however, suffer from long acquisition times (~100 sec) and cannot report on the spatial variation of micro-viscosities in a rapid imaging mode compatible with live cell imaging. Recently, fluorescent molecular rotors have emerged as novel, environmentally-sensitive probes capable of generating high-resolution images of the spatial distribution of micro-viscosities in a biological sample.[10–13] However, the current molecular rotor approach has low organelle specificity and does not allow for protein-specific micro-environment measurements. This missing protein-specific information, if obtainable, would enhance our understanding of the dynamics and function of proteins inside cells. The fluorescence properties (intensity, lifetime and quantum yield) of genetically-encoded fluorescent proteins (such as GFP) are, unfortunately, insensitive to the medium viscosity because of shielding and isolation of the chromophore from the surroundings by the protein $\beta$-barrels.[14] In addition, fluorescence anisotropy cannot report on the viscosity

experienced by stationary proteins (e.g., H2B) inside cells. In the present study, we explored the emerging chemical tagging technology and developed a hybrid genetic-chemical eDHFR-TMP-Cy3 rotor tag to report protein-specific micro-viscosity by fluorescence lifetime imaging microscopy (FLIM).

### 3.3 Design and synthesis of TMP-Cy3 probe

The emerging chemical tagging technology has offered a route to selectively label a protein of interest *in vivo* with an organic fluorescent dye in a genetically encoded manner.[15] The TMP-tag, which is designed around the nano-molar affinity (and recent covalent) interaction between *E. coli* dihydrofolate reductase (eDHFR) and trimethoprim (TMP), stands out as one of the few tags that work inside living cells with a high labeling specificity and efficiency.[16] The DNA sequence that encodes the small (18 kDa) and monomeric eDHFR is genetically fused to a protein of interest, and then a highly cell-permeable TMP-dye conjugate is introduced. The TMP-dye conjugate diffuses into the cell and recognizes the eDHFR fusion protein. By incorporating bright organic fluorophores such as Atto dyes, the TMP-tag has demonstrated its utility in super-resolution microscopy and in single-molecule biophysics.[17, 18]

We selected Cy3 as the conjugation probe because of its environment-sensitive fluorescence lifetime. The photophysical properties of Cy3 have been well studied.[19–22] As shown in Figure 3-1a, after excitation, in addition to the radiative decay pathway generating fluorescence, Cy3 can also isomerize from the *trans*-to the *cis*- configuration through a torsional motion, bringing Cy3 back to its ground state without photon emission. As a result, the measurable fluorescence lifetime, $\tau$,

$$\tau(\eta) = \frac{1}{k_{sp} + k_{nr}(\eta) + k_{i.c.}}$$

is the inverse sum of the radiative decay rate constant of spontaneous emission $k_{sp}$, the non-radiative decay rate, $k_{nr}$, which is sensitive to the medium viscosity, $\eta$, and direct internal conversion from the excited state, $k_{i.c.}$.[23] In a low-viscosity environment such as in water, the non-radiative photo-isomerization pathway of Cy3 is dominant and the fluorescence lifetime is short (<0.2 ns). A viscous medium will hinder the rate of torsional motion on the potential surface of the excited state and prolong the fluorescence lifetime, as explained by the Kramers theory on barrier crossing.[24] Although cyanine derivatives have been widely used in applications from *in vitro* protein labeling to *in vivo* animal diagnostic imaging,[25] the environmentally-sensitive photo-physical properties of cyanine derivatives have not been explored until recently. Notably, in two recent *in vitro* single-molecule studies, Cy3 was exploited as a novel local reporter to probe real-time protein binding onto DNA.[26,27] These emerging results encourage us to harness the environmental sensitivity of Cy3 in live cells for imaging.

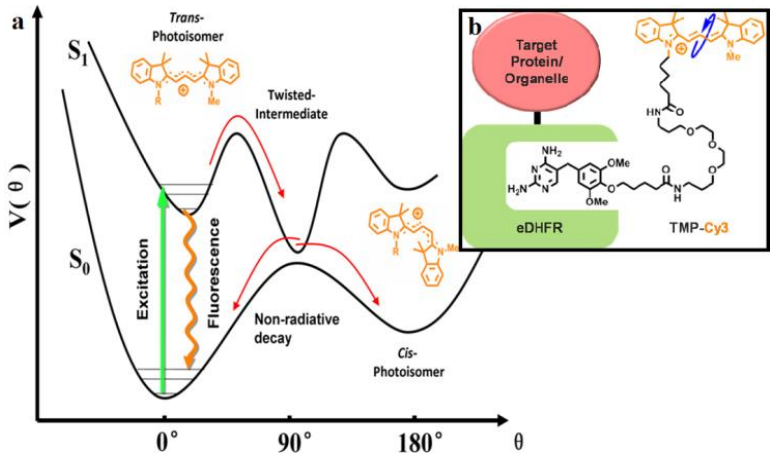



**Figure 3-1. Cy3 photophysics and TMP-Cy3 structure**. (a) Cy3 photophysics (b) Design of eDHFR-TMP-Cy3 probe

We developed a TMP–Cy3 probe for live cell imaging (Figure 3-1b). A cell-permeable TMP–Cy3 conjugate was synthesized by modularized conjugation of TMP–$NH_2$ with a sulfonate-free version of Cy3. A flexible polyethylene glycol (PEG) spacer was introduced between TMP and Cy3 to minimize the potential influence of the eDHFR protein on the nearby Cy3 probe.

### 3.4 *In vitro* characterization of TMP-Cy3 as an environmental sensor

TMP–Cy3 was tested as an environment sensor *in vitro* before its use in cellular experiments (Figure 3-2). Fluorescence intensity and lifetime measurements on TMP–Cy3 were carried out in glycerol–water solutions with different viscosities. All measurements were made on a home-built frequency-domain FLIM microscope. The fluorescence lifetime of TMP–Cy3 increases from 0.2 ± 0.1 ns to 1.6 ± 0.2 ns with increasing glycerol volume fraction from 20% to 100%, consistent with a restriction of torsional motion and a hindered non-radiative decay. The fluorescence of TMP–Cy3 and eDHFR–TMP–Cy3 was compared in order to examine the potential perturbation effect of the protein on the nearby fluorophore. As shown in Figure 3-2A, a minimal effect of the eDHFR binding on the lifetime and brightness of TMP–Cy3 was observed, most likely attributed to the long PEG spacer designed between TMP and Cy3.

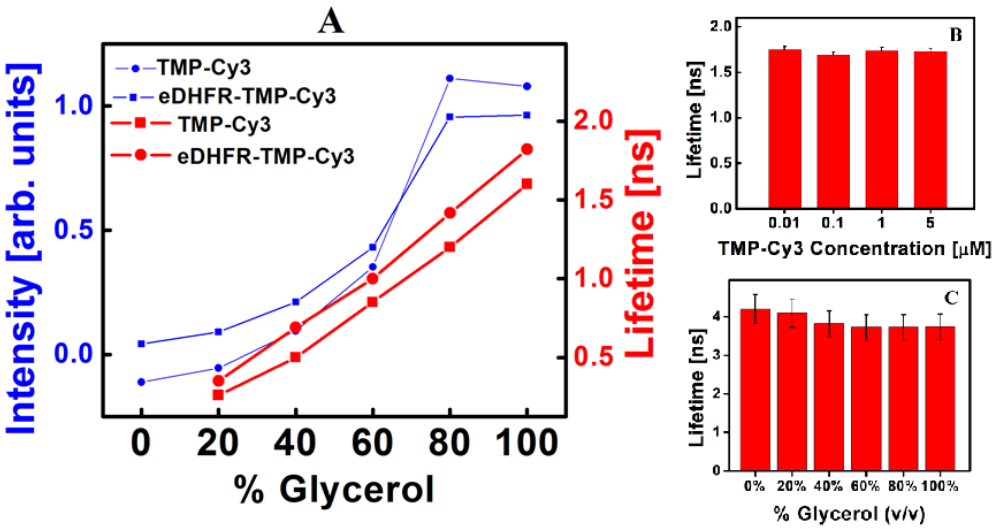



**Figure 3-2. TMP-Cy3 as an environmental sensor.** (A) Fluorescence intensity and lifetime of TMP-Cy3 and eDHFR-TMP-Cy3. (B) Concentration independence of TMP-Cy3, (C) Rhodamine insensitivity to viscosity.

For use as a specific micro-environment sensor inside cells, TMP–Cy3 should have negligible interactions with DNA, proteins and ions. Indeed, this is the case, as confirmed by *in vitro* spectroscopy experiments of TMP–Cy3 in solutions of varying DNA, bovine serum albumin (BSA) and NaCl concentrations (see Supporting Information section for details). TMP–Cy3, however, does weakly interact with micelles formed by SDS (Sodium Dodecyl Sulfate). The fluorescence lifetime of TMP–Cy3 is independent of the fluorophore concentration (Figure 3-2B), which is an advantage in quantitative FLIM studies as the concentration of the fluorophore within live cells is not well controlled. At concentrations from 10 nm to 5 μM, the fluorescence lifetime of TMP–Cy3 is constant (Figure 3-2B). Rhodamine 6G, a common non-rotor dye, has a constant fluorescence lifetime with varying glycerol concentrations (Figure 3-2C), underscoring the necessity of flexible rotor tags for viscosity imaging.

## 3.5 TMP-Cy3 as a live-cell protein-specific environmental sensor

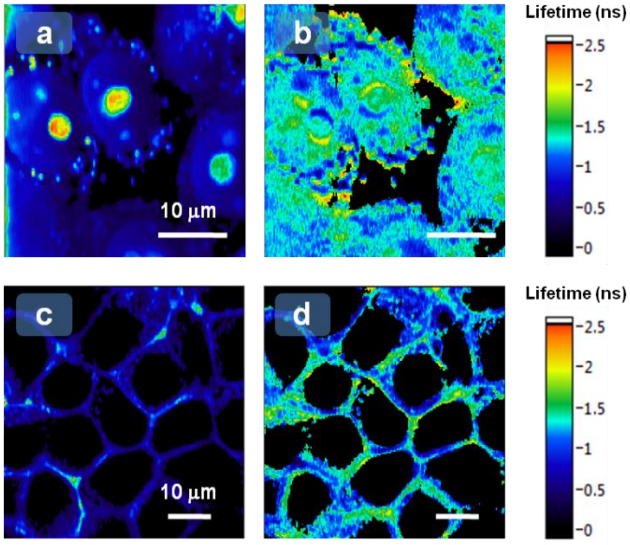



**Figure 3-3. TMP-Cy3 as protein specific environmental sensors in live cells.** Fluorescence intensity (a, c) and fluorescence lifetime (b, d) images of H2B-eDHFR-TMP-Cy3 (a, b) and PMLS-eDHFR-TMP-Cy3 (c, d) in live HEK cells

Encouraged by the *in vitro* experiments, we moved on to cellular experiments to evaluate the ability of TMP–Cy3 to map local micro-environments. To target the molecular rotor to the cell nucleus, eDHFR was fused to histone H2B protein. We also studied the cytosol environment in the vicinity of the cell membrane by anchoring eDHFR with a plasma membrane localization signal (PMLS). We transiently transfected HEK 293T cells with plasmids encoding H2B-eDHFR or PMLS-eDHFR fusion protein, then incubated the cells with our synthesized, cell-permeable TMP–Cy3 for 10 min, and imaged them. The H2B-eDHFR images show minor degrees of non-specific granular-shaped binding of TMP–Cy3 to lipid-rich organelles.[28] Nevertheless, oval-shaped nuclei with distinct nucleoli were clearly observed (Figure 3-3a). Characteristic plasma membrane patterns were observed in cells expressing PMLS-eDHFR (Figure 3-3c). Therefore, this approach would be suitable

for monitoring the nuclear and plasma membrane environment as non-specific staining was minor inside the nucleus and near the plasma membrane.

FLIM microscopy was used to image the distributions of local environments experienced by H2B within the cell nuclei and cytosol viscosities near the cell membrane. FLIM imaging (Figure 3b) revealed that the nuclear viscosity is high, comparable to glycerol solutions from 50% volume fraction ($\sim$1 ns lifetime) up to considerably higher values in certain regions (>2.0 ns lifetime). The average lifetime over the two brightest nuclei with clearly visible nucleoli to the left-of-center in Figure 3-3a is 1.4 $\pm$ 0.3 ns, indicative of a viscous and crowded environment. Our lifetime measurements correspond to average viscosities of around 60–70 cP within the nucleus, comparable to other reports.[13] There is also remarkable heterogeneity in the distributions of the measured micro-environment, offering information not available from the confocal fluorescence intensity image. This heterogeneity may be related to the recently reported heterogeneous level of chromatin compaction detected by fluorescence anisotropy imaging of H2B-EGFP.[29] FLIM images were also captured for TMP–Cy3 labeled PMLS-eDHFR in the vicinity of cell membranes (Figure 3-3d). In this scenario, the average lifetime is typically 0.9 $\pm$ 0.2 ns, indicating a relatively less viscous and crowded area of the cell cytoplasm compared to the nucleus experienced by H2B.

### 3.6 Conclusion

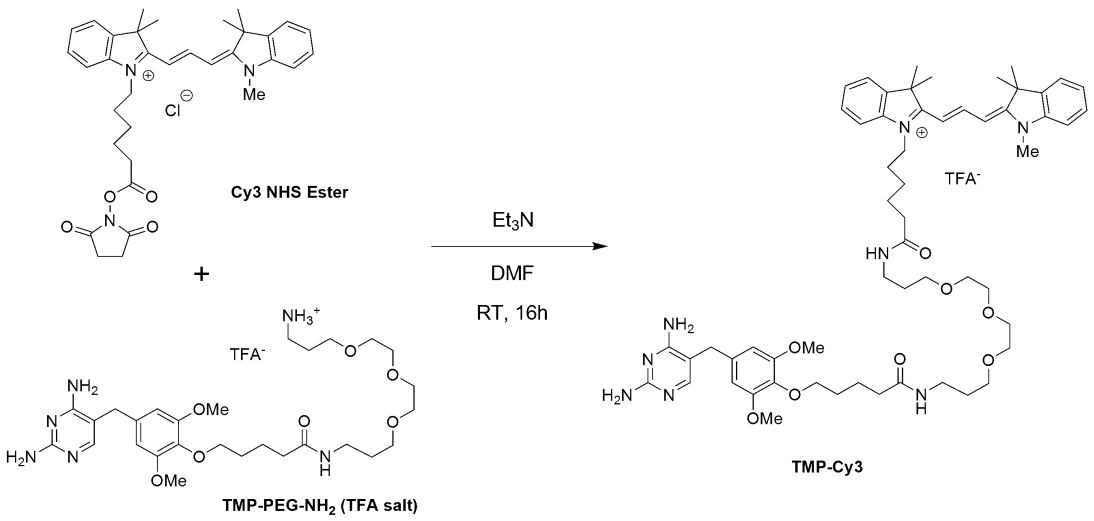
We developed a hybrid genetic-chemical molecular rotor tag (eDHFR–TMP–Cy3) to measure protein-specific local environments in live cells using FLIM. Although we only used the TMP-tag, the same Cy3 rotor moiety can be readily applied to other chemical tagging techniques such as the SNAP, CLIP, and HaloTag.[15,30] This methodology, with its

good genetically-encoded specificity, high spatial-temporal resolution and simple interpretation, could provide valuable mechanistic information about protein function in the complex and constantly changing cellular environment. For example, the observed heterogeneous micro-environment could have broader implications in understanding chromatin condensation and transcription control within live cells.[31] Chemical tags can be engineered with arbitrary open-structure biophysical probes that are exposed to their surroundings and can sense the local environment more sensitively than regular fluorescent proteins. To our knowledge, this advantageous aspect of chemical tags has been largely unexplored. This protein/organelle specific FLIM technique should be useful for evaluating a wider variety of protein or organelle-specific cellular micro-environments.

## 3.7 Supporting information and experimental methods

### 3.7.1 Synthesis and characterization of TMP-Cy3

Anhydrous dimethylformamide was obtained from Aldrich, triethylamine was obtained from Fluka. Cy3 NHS ester was obtained from Lumiprobe, LLC. Nuclear magnetic resonance (NMR) spectra were recorded on a Bruker 400 (400MHz) Fourier Transform (FT) NMR spectrometer at the Columbia University Chemistry Department. $^{1}$H NMR spectra are tabulated in the following order: multiplicity (s, singlet; d, doublet; t, triplet; m, multiplet). Electron spray ionization (ESI) MS were recorded on a JMS-LC Mate mass spectrometer.




**Scheme 3-S1. Synthesis of TMP-Cy3**

In a 5.0 mL vial TMP-PEG-$NH_2$ (TFA salt) (1.0 mg, 1.4 μmol) and Cy3 NHS-ester (1.2 mg, 2.0 μmol) were dissolved in 0.2 mL of anhydrous dimethylformamide (DMF). Triethylamine (TEA) (5.0 μL, 50.0 μmol) was added to the vial and the reaction was stirred overnight at low light and at room temperature. The reaction mixture was concentrated to dryness, re-dissolved in 3.0 mL of $H_2O/CH_3CN$ 4:1 (v/v) and purified by preparative HPLC using a linear gradient of solvent $H_2O$ (0.1% TFA) / $CH_3CN$ 80 / 20 to 40 / 60 over 40 min to give 1.6 mg TMP-Cy3 as a pink-brown solid (near quantitative) with a retention time of 24.6 min.

**$^1$H NMR** (400 MHz, $CD_3OD$) δ ppm: 8.55 (t, *J* = 13.6 Hz, 1 H); 7.55 (d, *J* = 7.2 Hz, 2 H); 7.45 (m, 2 H); 7.37-7.29 (m, 4 H); 7.22 (s, 1 H); 6.56 (s, 2 H); 6.45 (d, *J* = 13.6 Hz, 1 H); 6.43 (d, *J* = 13.6 Hz, 1 H); 4.15 (t, *J* = 7.0 Hz, 2 H); 3.91 (t, *J* = 6.2 Hz, 2 H); 3.79 (s, 6 H); 3.69 (s, 3 H); 3.66 (s, 2 H); 3.62-3.60 (m, 4 H); 3.57-3.54 (m, 4 H); 3.50 (t, *J* = 6.2 Hz, 2 H); 3.47 (t, *J* = 6.4 Hz, 2 H); 3.25 (t, *J* = 6.8 Hz, 2 H); 3.22 (t, *J* = 6.8 Hz, 2 H); 2.25 (t, *J* = 7.2 Hz, 2 H); 2.21 (t, *J* = 7.2 Hz, 2 H); 1.88-1.68 (m, 12 H); 1.77 (s, 12 H); 1.48 (m, 2 H). MS (ESI+) m/z Calcd for $C_{58}H_{81}O_8N_8^+$ $[M]^+$: 1017.62, Found: 1017.93

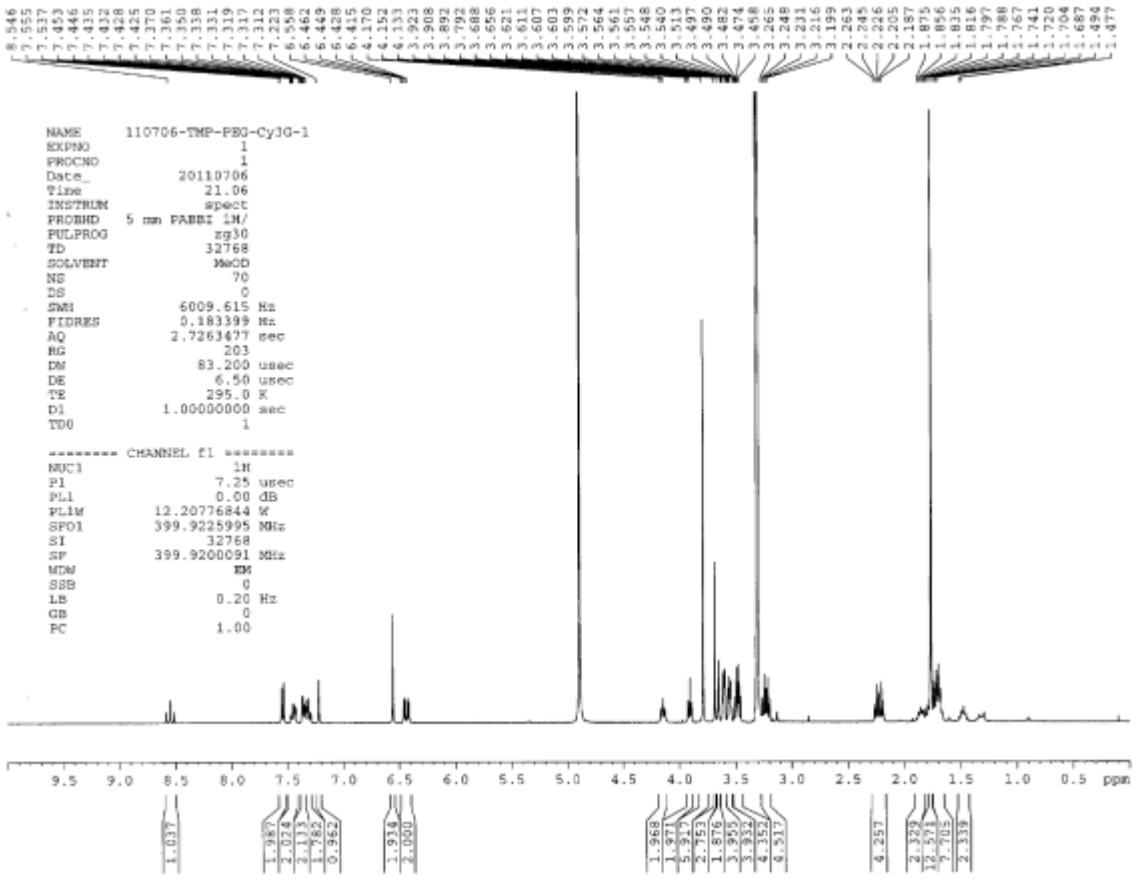



**Figure 3-S1.$^{1}$H NMR spectrum of TMP-Cy3 ($CD_3OD$, 400 MHz)**.

### 3.7.2 Cell culture, transfection and labeling

HEK 293T cells were cultured in DMEM w/ glutamine (Gibco #11995) with 10% v/v fetal bovine serum and 1% v/v Pen./Strep. Cells were maintained under 5% $CO_2$ at 37 ℃. For *in vivo* labelling, cells were plated in eight-well chambered #1 borosilicate coverglass-bottomed chambers (Thermo, Nunc 155411) twenty-four hours before transfection with expression plasmids for H2B-eDHFR or PMLS-eDHFR (0.4 μg DNA for one well) using Fugene HD (Roche). Twenty four hours after transfection, 300 μL of cell culture media was

added and 0.24 μL TMP-Cy3 stock solution (1.25 mM in DMF) for a final TMP-Cy3 concentration of 1 μM. Cells were incubated with this staining solution for 10 min at 37°C, and then washed twice with fresh media before imaging at room temperature (25℃). All spectroscopy and imaging experiments were performed at room temperature.

### 3.7.3 Fluorescence Imaging

Imaging was performed using a home-built confocal microscope configured on an inverted microscope body (IX71, Olympus) and operated in a sample scanning mode. A 532 nm diode laser (Compass 215M, Coherent) was used to excite the Cy3-labeled cells. A number of confocal fluorescence images were also captured on a commercial confocal laser scanning microscope (SP5, Leica) with the 488 and 514 nm lines of the Argon Ion laser for Fluoroscein and Cy3, respectively, and a 594nm laser line for mCherry imaging. All measurements were performed at room temperature.

Untransfected cells (Figure 3-S2, top row) display minor degrees of background staining but no localization of TMP-Cy3 to the nuclei as in the middle row of Figure 3-S2, where TMP-Cy3 rapidly localizes to H2B-eDHFR in the nucleus of H2B-eDHFR transfected 293T cells. The nuclei are quite dark in the first row of Figure 3-S2, compared to the row below, where TMPCy3 targeted H2B-eDHFR labeled protein within the nuclei.

Several tests were performed to assess whether H2B-eDHFR-TMP-Cy3 fusion protein distributes the same as native H2B protein. Cells were also transfected to express H2B-mCherry, mCherry is a bright red fluorescent protein that provides direct imaging of H2B in live cells. We imaged H2B-mCherry transfected cells simultaneously with H2B-eDHFR transfected cells in order to verify that eDHFR does indeed target H2B protein with high specificity in live cells. Cells were co-labeled with H2B-eDHFR-TMP-Cy3 and H2B-

mCherry and the degree of overlap was assessed (Figure 3-S2). There is good overlap between H2B-eDHFR-TMP-Cy3 staining and H2B-mCherry images at the bottom right of Figure 3-S2. Confocal fluorescence images are overlaid with DIC in the last column to the right. From left to right all images present confocal fluorescence from Cy3, mCherry, the DIC image and the overlay of fluorescence with DIC.

In order to address the non-specific staining of TMP-Cy3 to mitochondria and potentially other cellular organelles, we imaged 293T cells transfected with H2B-eDHFR and then stained with TMP-Fluorescein, Figure 3-S3. In this control image we observe that eDHFR is well targeted to cell nuclei and doesn't perturb cell functions such as chromatin formation. In addition TMP-Fluorescein did not exhibit the background staining that Cy3 has shown, for example in Figure 3-2, where H2B-TMP-Cy3 exhibits a markedly higher level of background staining compared to H2B-mCherry. However, Cy3 does exhibit viscosity sensitivity, unlike Fluorescein. Efforts are underway towards developing a better-behaved molecular rotor, and other environmental probes.

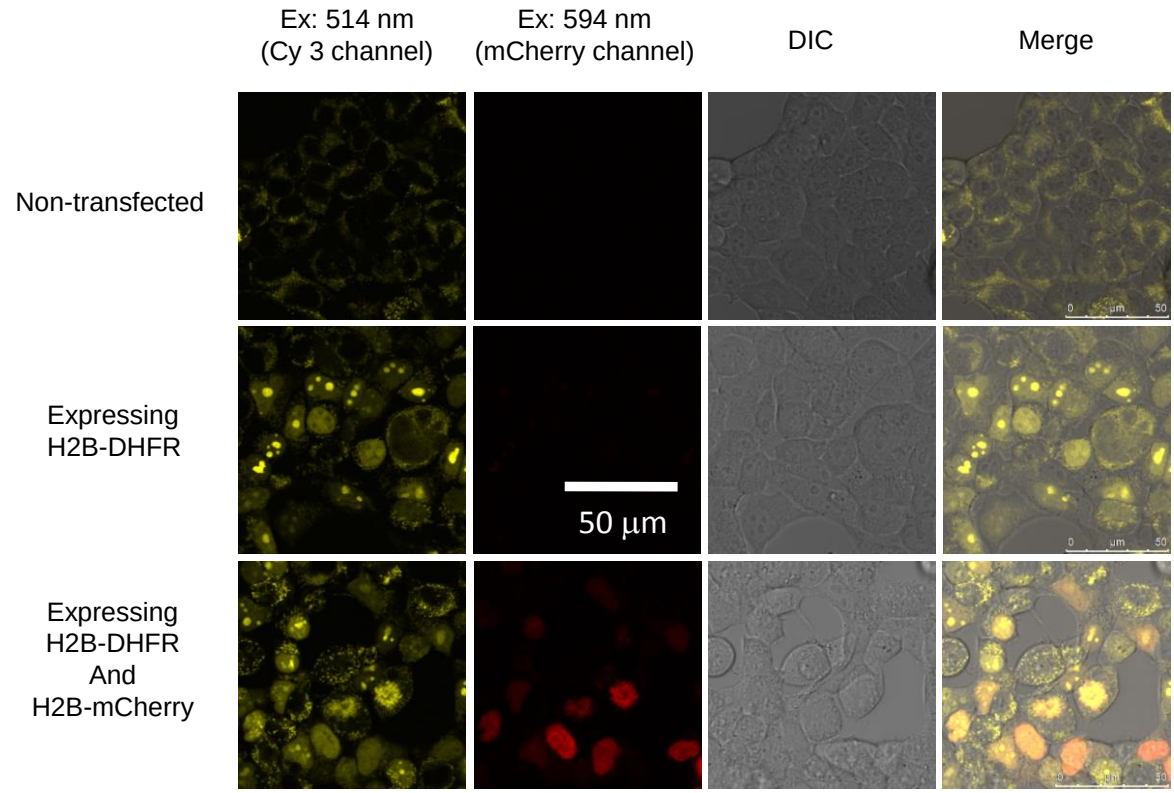



**Figure 3-S2. Confocal fluorescence of untransfected (top row), H2B-eDHFR transfected (middle row) and H2B-eDHFR and H2B-mCherry dual transfected cells (bottom row).** Fluorescence was monitored from left to right with 514nm excitation for Cy3 and 594nm excitation for H2B-mCherry, the DIC image is provided in the third column and the overlay of the Cy3, mCherry fluorescence and DIC is shown in the last column.

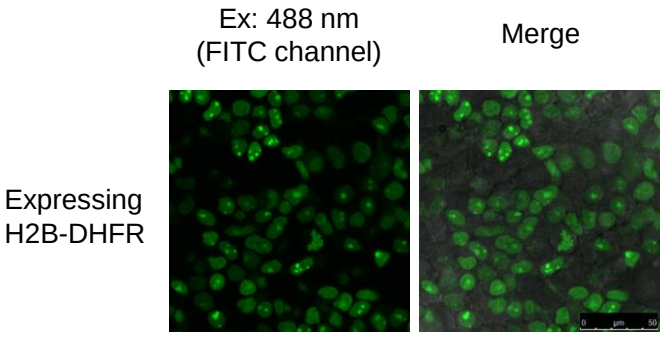



**Figure 3-S3. Confocal fluorescence of 293T cells transfected with H2B-eDHFR and labeled with Fluorescein.** Cells were excited with a 488nm laser line and confocal fluorescence images were acquired. Scale bar 50 µm.

### 3.7.4 Fluorescence lifetime imaging microscopy (FLIM)

A home-built FLIM microscope capability was configured by modulating the 532nm excitation laser at 50 MHz with an acousto-optic modulator (3200-121, Crystal Technology) driven by a sine wave from an arbitrary function generator (AFG 3102, Tektronix). The 532 nm beam was focused by a 50 mm achromatic lens into the AOM, which was centered in a 4X beam expanding telescope. The undiffracted zeroth order beam was selected. A dichroic filter (Di01-R532-25x35, Semrock) directed the sinusoidally-modulated excitation laser onto the sample and the same objective was used to collect the sinusoidally-modulated and phase-shifted fluorescence. A high-NA water lens (60x 1.2NA UPLANSAPO, Olympus) was used for all FLIM images. A Hamamatsu R9110 photomultiplier tube (PMT) was placed at the microscope side port primary image plane, with a 50 μm pinhole serving as the confocal pinhole. After passing through the microscope internal optics and a long-pass filter, the fluorescence was detected by the PMT. The PMT electrical signal was pre-amplified by a wide-bandwidth (up to 50 MHz) pre-amplifier (C6438-01, Hamamatsu) and input directly into a dual-phase RF lock-in amplifier (SR844, Stanford Research Systems).

A frequency-domain fluorescence lifetime approach was used with sinusoidal modulation at 50 MHz and subsequent de-modulation and extraction of the phase lag using a dual-phase wide bandwidth lock-in amplifier. Both the in-phase **(X)** and out-of-phase component **(Y)** of the sinusoidally-modulated fluorescence, relative to the excitation sine wave, are determined by the lock-in amplifier and the phase difference (lag) is computed and defined as:

$$\theta \equiv \arctan(Y/X) \qquad \textbf{(Equation 3-S1)}$$

The phase lag was measured for each pixel and recorded using home-written software in LabView that also moved the nanometric sample-scanning stage. Pixel dwell times from 100 μs to 3ms were used. Pixel dwell times were set equal to the lock-in time constant. A lock-in time constant of 1 ms gave the best results (26 seconds per frame at 100x100 pixels), and a 300 μs time constant provided satisfactory imaging at a significantly faster rate (6 s per frame). For imaging, depending on the specific protein that was transfected either 3 volts were sent to the nanometric sample scanning stage, which defined a total scan area of 21 μm x 21 μm (typically used for the nuclear H2B protein images) or 9-10 Volts for a total scan area of 63-70 μm x 63-70 μm (typically used for the PMLS plasma membrane protein images). 100x100 or 200x200 pixels defined the imaging area for most images, with each pixel corresponding to 210 nm or 315 nm, both below or comparable to the diffraction limit of the objective lens used.

A modulation frequency of 50 MHz was used for frequency domain fluorescence lifetime measurements. This frequency was chosen because a) it is sufficient for frequency-domain phase fluorometry (~20ns per period) and b) it is an upper limit of performance for our AOM. At modulator driving frequencies above 50MHz there is a substantial decrease in modulation depth and a distortion of the laser output. Modulation depth and AOM stability are key parameters for the quality of the lifetime imaging. Several control experiments revealed that the phase shift of Cy3 in various glycerol mixtures was always between that of the laser response and Rhodamine 6G, Rhodamine 6G displaying the largest phase lag (most negative phase), Cy3 an intermediate lag, and the laser response always the most positive of the three. The phase of the sinusoidally-modulated fluorescence was compared to that of the excitation laser for all measurements. To determine the instrument response, the long-pass

filter was removed from its position right before the confocal pinhole and an image was taken of the modulated laser reflection from the glass coverslip surface, a mirrored microscope slide, or by focusing on the bottom of the 8 well coverglass-bottomed chambered cell slide. Care must be taken with this approach, as the focus may alter the phase shift. Focusing into a solution of a well-known dye like Rhodamine6G is more appropriate in some cases, as the distribution of phases is narrower and the solution presumably where the cells would be (minimizing optical phase shifts). The reflection image serves as a measurement of all the delays due to system electronics and optics. The reflected laser image is then averaged to give an instrumental phase. This instrumental phase must be subtracted from the total phase of the sample and instrument in order to determine the sample's phase lifetime. This phase difference between the fluorescence phase due to the sample and from the instrument, when referenced to the instrument without any sample, will give the sample's phase lifetime according to Equation 3-S2 below. After an image of the laser response, the long-pass filter is put back in and a phase is recorded for each pixel. There is sufficient leak-through of the 532nm laser light through the dichroic that only the long-pass filter right before the PMT need be removed to measure the instrument response. Several laser scans were taken at the beginning, middle, and end of an imaging session to verify that lifetime values were fairly consistent. The phase lag due to the sample is the phase recorded from the sample after subtracting the laser phase. Usually this is done by subtracting an average of points over a smooth laser image from the actual cell image. The fluorescence lifetime at each pixel was determined from the phase lag at each pixel according to the simple relation:

$$\tau_\varphi = \omega^{-1}\tan(\Delta\varphi) \qquad \textbf{(Equation 3-S2)}$$

where **ω** is the angular modulation frequency and **Δφ** is the experimentally observed phase lag. The sample was scanned with a piezo-electronic nanometric sample scanner (Physik Instrumente, PI nano), although the viscosity and chemical micro-environment is measured with a spatial resolution limited by that of the diffraction-limited optics (~300nm). All imaging and analysis was done on home-written software using LabView.

Fluorescence lifetime data of TMP-Cy3 in glycerol/water solutions of varying volume fraction from 0-100% were also taken on a single-photon counting fluorescence lifetime spectrometer (OB920, Edinburgh Analytical Instruments, U.K.) equipped with a PicoQuant 496nm pulsed LED diode as the excitation light source. The results were similar to what we obtained with the home-built FLIM microscope apparatus, for example, the measured fluorescence lifetime of TMP-Cy3 was the same for both (~0.3 ns).

### 3.7.5 Additional Spectroscopic Measurements on TMP-Cy3

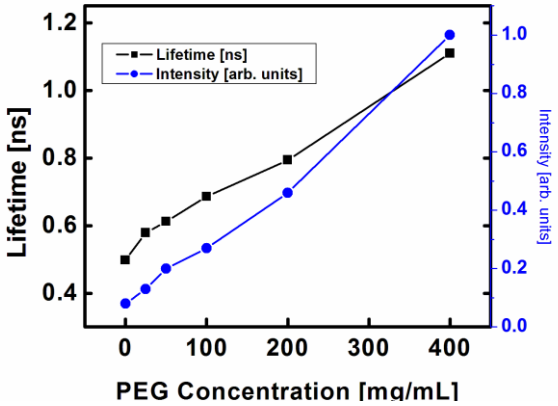



**Figure 3-S4. Fluorescence lifetime and intensity of TMP-Cy3 as a function of PEG concentration.**

As a further test of the effect of viscosity and molecular crowding on the lifetime and intensity of Cy3, both the lifetime and intensity were measured in various concentration solutions of polyethylene glycol (PEG) from 0, 25, 50, 100, 200, to 400 mg/mL of PEG. An

increase in both the lifetime and intensity were seen, with values approaching that of a 40-60% glycerol solution as the concentration was increased from 200 to 400 mg/mL.

As opposed to simple tests in glycerol and PEG solutions, using Cy3 to report on local cellular micro-viscosities requires further careful tests on solutions approximating the cellular environment. Two important concerns in the confocal intensity and FLIM mapping of intracellular viscosity are non-specific binding to cellular macromolecules (DNA, protein, lipid membranes, etc.) and the effect of other parameters such as polarity of the environment on the fluorescence intensity and lifetime of Cy3. Therefore, we performed systematic fluorescence intensity measurements in several solutions (DCM, ethanol, DMSO, dioxane, SDS and NaCl) and biomolecules (DNA and proteins). Solutions of 1 μM TMP-Cy3 were prepared for all measurements and the absorption and fluorescence of Cy3 were recorded with a commercial plate reader (Tecan Infinite 200). Error was calculated by comparing the differences between several measurements on similar samples in the same run. All fluorescence intensities were normalized with respect to TMP-Cy3 in $H_2O$. Fluorescence intensity measurements can be surrogates for lifetime measurements as they are affected by the same non-radiative decay mechanisms, although there are important differences as seen with the BSA lifetime data in Figure 3-S11.

We tested the amphiphilic detergent Sodium Dodecyl Sulfate (SDS). SDS has an effect on the fluorescence intensity of TMP-Cy3 as observed by an increase in the fluorescence intensity by ~80% between pure $H_2O$ and a 10 mg/mL solution of SDS. Apparently there is a non-specific interaction between TMP-Cy3 and, presumably, the micelles formed by SDS. Interestingly, there is no significant change in the fluorescence intensity of TMP-Cy3 in going from a 10 mg/mL to a 20 mg/mL solution of SDS. The

observed increase in TMP-Cy3 fluorescence intensity in solutions of SDS may help explain the non-specific binding of Cy3 to mitochondria and lipid membranes in our cellular measurements.

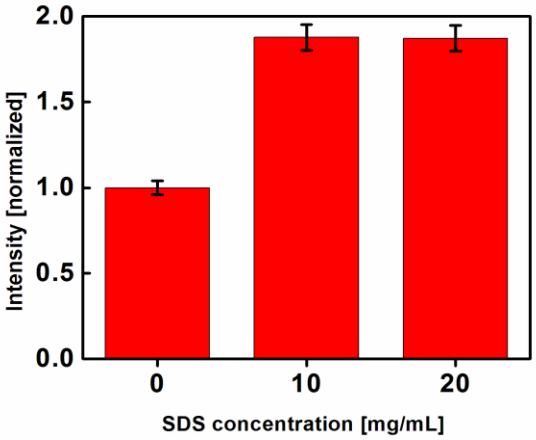



**Figure 3-S5. Fluorescence intensity of TMP-Cy3 as a function of SDS concentration.** There is a significant increase in the fluorescence intensity of a 1 μM TMP-Cy3 between pure $H_2O$ and 10 mg/mL of added SDS detergent, but not at higher concentration.

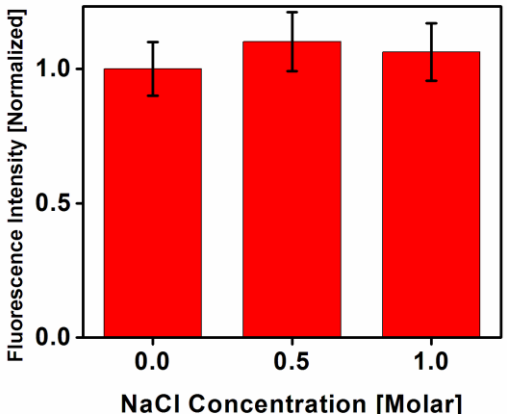



**Figure 3-S6. Fluorescence intensity of TMP-Cy3 as a function of NaCl concentration.** There is a minimal effect of ionic concentration on the fluorescence of TMP-Cy3.

We performed assays of the effect of ionic strength on TMP-Cy3 fluorescence in solution. As shown in Figure 3-S6 below, there is a minimal effect on the fluorescence

intensity of TMP-Cy3 with increasing concentration of NaCl from 0 M to a 1.0 M concentration solution of $H_2O$/NaCl.

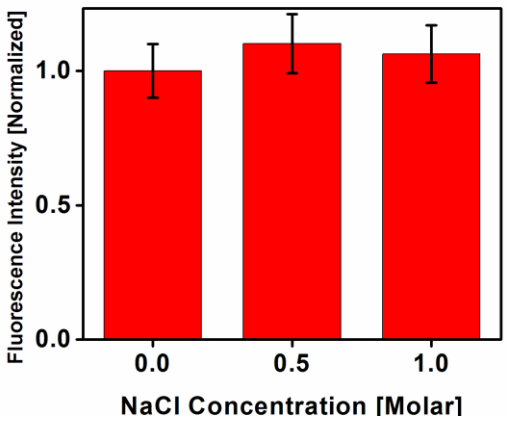



**Figure 3-S6. Fluorescence intensity of TMP-Cy3 as a function of NaCl concentration.** There is a minimal effect of ionic concentration on the fluorescence of TMP-Cy3.

We performed several assays of the effect of the presence of higher and higher concentrations of DNA in solution on the fluorescence intensity of TMP-Cy3 in solution. As shown in Figure 3-S7 below, there is no significant change in the fluorescence intensity upon addition of increasing amounts of DNA.

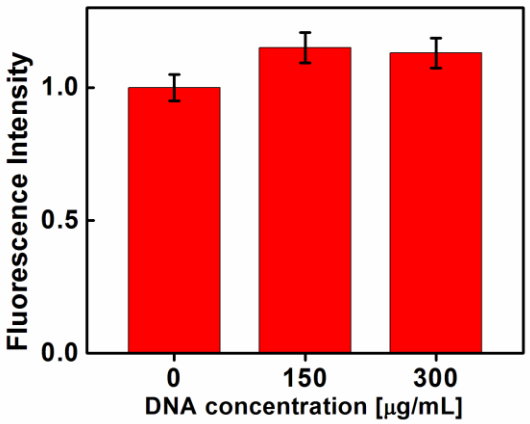



**Figure 3-S7. Fluorescence intensity of 1 µM TMP-Cy3 as a function of DNA concentration**. normalized with respect to 1 µM TMP-Cy3 in $H_2O$.

Although not present in the cell, we tested the fluorescence of TMP-Cy3 in a variety of solvents as follows: DMSO, Ethanol, DCM, and in an assay of Dioxane/H2O with varying volume percentages. In pure solvents, there is a notable change in the fluorescence intensity of TMP-Cy3, as shown in Figure 3-S8.

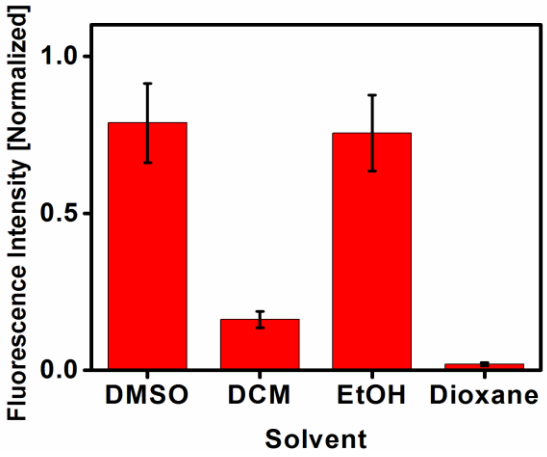



**Figure 3-S8. Solvent effect of TMP-Cy3 fluorescence intensity**.

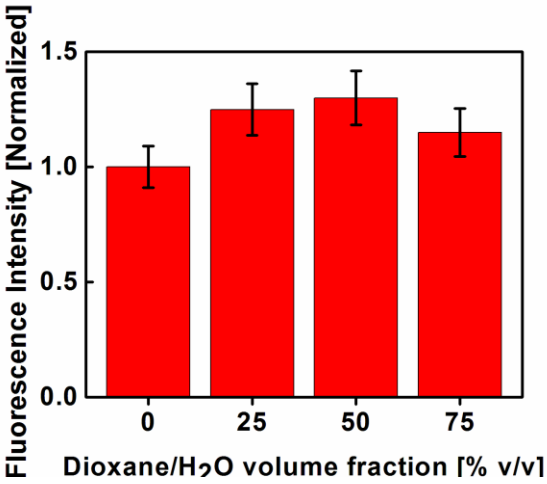



**Figure 3-S9. Dioxane/$H_2O$ solvent effect of TMP-Cy3 fluorescence intensity.** There is an increase in the fluorescence intensity of TMP-Cy3 with increasing dioxane percentage, by about 30% from 0% to 50%, followed by a slight decrease.

In solutions of Dioxane/$H_2O$ from 0% to 75% volume/volume of Dioxane, there is an increase in the fluorescence intensity in going from 0% to 50% dioxane by a factor of

about 30% (Figure 3-S9). However, in a pure Dioxane solution, the fluorescence intensity drops considerably, to <10% of that in pure $H_2O$. Error bars for all charts were calculated from the error between several identical measurements.

We also had to assess the effect of typical chemical species that can be found in the intracellular environment on the fluorescence lifetime, since we used the fluorescence lifetime for viscosity measurements. Additional experiments were performed in solutions of NaCl, SDS (where there was a significant change in the fluorescence intensity, but not the lifetime), the protein BSA, and other chemical species. No significant change was seen in the fluorescence lifetime upon changing the concentration of a series of some of the types of chemical species that would be encountered in the cell (ions, lipid domains, protein). A commercial lifetime fluorometer (OB920, Edinburgh Analytical Instruments, U.K.) with a PicoQuant 496nm pulsed LED diode as the excitation source was also used for the solution lifetime measurements, and we obtained values comparable to our homebuilt FLIM apparatus. Each run took 20 minutes – which would not be compatible with biological imaging. We do note that there was evidence of multi-exponential behavior in the BSA lifetimes, with one component around 50% at ~1.5ns and another component around ~50% closer to 3ns. Multi-exponential components can also be resolved with frequency domain FLIM, but would require an Electro-Optic Modulator and a larger frequency sweep or more expensive time-correlated single-photon counting (TCSPC) electronics. The data for water, NaCl, SDS, and BSA are shown below in Table 3-S1.

**Table 3-S1. Raw lifetime data from a commercial lifetime fluorometer for water, SDS, NaCl, and BSA solutions of TMP-Cy3.** There is an observed high lifetime for concentrated solutions of BSA, but lifetimes are more constant than intensities.

| |
|---|
| H2O: 0.30 ns<br><br>SDS 10 mg/ml : 0.51 ns<br>SDS 20 mg/ml : 0.50 ns<br><br>NaCl 0.5 M : 0.33 ns<br>NaCl 1.0 M: 0.34 ns<br><br>BSA 50 mg/ml : 2.22 ns (if double-exponential: 35% 1.49 ns and 65% 2.76 ns )<br>BSA 100 mg/ml : 2.23 ns (if fit in double-exponential: 49% 1.62 ns and 51% 3.00 ns)<br>BSA 200 mg/ml : 2.18 ns (if fit in double-exponential: 47% 1.55 ns and 53% 2.93 ns) |

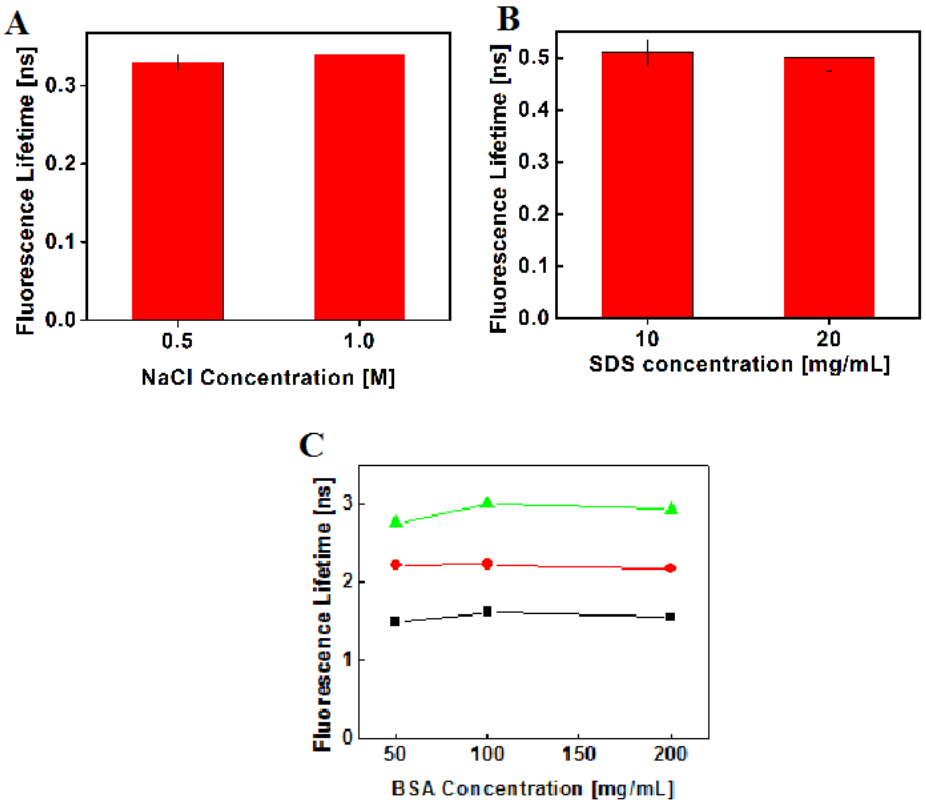



**Figure 3-S10. Fluorescence lifetime of TMP-Cy3 in the presence of chemicals**. (A) There is no observable change in the fluorescence lifetime with increasing NaCl from 0.5 to 1M or (B) SDS concentration from 10 to 20 mg/mL. (C) BSA lifetimes are fairly constant, although high, and there is the possibility of multi-exponential behavior with a lower-lifetime component (black), average (red), and high-lifetime component (green).

With BSA, there is an increasing fluorescence lifetime with increasing BSA concentration; this indicates that the fluorescence lifetime is indeed reporting on local viscosity as increasing protein concentration can dramatically increase viscosity. The fluorescence lifetime can also reveal information about protein environments that is obscured or less obvious from the fluorescence intensity alone. Overall, the fluorescence lifetime data is more constant than the intensity data but there is a demonstrable protein effect on lifetimes not seen in the fluorescence intensity data alone.

### 3.8 References


1. K. E. Handwerger, J. G. Gall, *Trends Cell Biol*. 2006, **16,** 19–26.
2. I. Lang, M. Scholz, R. Peters, J. *Cell Biol*. 1986, **102,** 1183–1190.
3. W. F. Marshall, A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, J. W. Sedat, *Curr. Biol*.1997, **7,** 930–939.
4. T. Kues, A. Dickmanns, R. Lührmann, R. Peters, U. Kubitscheck, *Proc. Natl. Acad. Sci. USA*. 2001, **98,** 12021–12026.
5. K. Luby-Phelps, *Int. Rev. Cytol*. 2000, **192,** 189–221.
6. J. W. Wojcieszyn, R. A. Schlegel, E. S. Wu, K. A. Jacobson, *Proc. Natl. Acad. Sci. USA* 1981, **78,** 4407-4410.
7. L. Liang, X. Wang, D. Xing, T. Chen, W. R. Chen, *J. Biomed. Opt*. 2009, **14,** 024013-1-9.
8. O. Seksek, J. Biwersi, A. S. Verkman, *J. Cell Biol*. 1997, **138,** 131-142.
9. D. Grünwald, R. M. Martin, V. Buschmann, D. P. Bazett-Jones, H. Leonhardt, U. Kubitscheck, M. C. Cardoso, *Biophys. J*. 2008, **94,** 2847–2858.
10. M. A. Haidekker, T. P. Brady, D. Lichlyter, E. A. Theodorakis, *J. Am. Chem. Soc.*, 2006, **128,** 398-399.
11. M. K. Kuimova, S. W. Botchway, A. W. Parker, M. Balaz, H. A. Collins, H. L. Anderson, K. Suhling, P. R. Ogilby, *Nature Chem*. 2009, **1,** 69-73.
12. X. Peng, Z. Yang, J. Wang, J. Fan, Y. He, F. Song, B. Wang, S. Sun, J. Qu, J. Qi, M. Yan, *J. Am. Chem. Soc*. 2011, **133,** 6626-6635.
13. M. K. Kuimova, G. Yahioglu, J. A. Levitt, K. Suhling, *J. Am. Chem. Soc.* 2008, **130,** 6672-6673.

14. K. Suhling, J. Siegel, D.Phillips, P. M.French, S. Lévêque-Fort, S. E.Webb, D. M. Davis, *Biophys. J*. 2002, **83,** 3589–3595.

15. C. Jing, V. W. Cornish, *Acc. Chem. Res*. 2011, **44,** 784-792.

16. L. W. Miller, Y. Cai, M. P. Sheetz, V. W.Cornish, *Nat. Methods* 2005**, 2,** 255-257.

17. R. Wombacher, M. Heidbreder, S. van de Linde, M. P. Sheetz, M. Heilemann, V. Cornish, M. Sauer, *Nat. Methods*, 2010, **7,** 717-719.

18. A. A. Hoskins, L. J. Friedman, S. S. Gallagher, D. J. Crawford, E. G. Anderson, R.Wombacher, N. Ramirez, V. W. Cornish, J. Gelles, M. J. Moore, *Science*, 2011, **331,** 1289-1295.

19. P. F. Aramendia, R. M. Negri, E. S. Roman, *J. Phys. Chem*., 1994, **98,** 3165–3173.

20. M. E. Sanborn, B. K. Connolly, K. Gurunathan, M. Levitus, *J. Phys. Chem*. B 2007, **111,** 11064-11074.

21. M. Levitus, S. Ranjit, *Q. Rev. Biophys*. 2011, **44,** 123-151.

22. H. S. Muddana, T. T. Morgan, J. H. Adair, P. J. Butler, *Nano Lett*., 2009, **9,** 1559-1566.

23. M. Y. Berezin, S. Achilefu, *Chem. Reviews*, 2010, **10,** 2641-2684.

24. H. A. Kramers, *Physica*, 1941, **7,** 284–304.

25. A. Samanta et al. *Angew. Chem. Int. Ed*. 2011, **50,** 6089-6092.

26. G. Luo, M. Wang, W. H. Konigsberg, X. S. Xie, *Proc. Natl. Acad. Sci. USA.*, 2007, **104,** 12610-12615.

27. S. Myong, S. Cui, V. P. Cornish, A. Kirchhofer, M. U. Gack, J. U. Jung, K. P. Hopfner, T. Ha, *Science*. 2009, **323,** 1070-1074.

28. L. V. Johnson, M. L. Walsh, B. J. Bockus, L. B. Chen, *J. Cell. Biol*. 1981, **88,** 526-535.

29. B. Banerjee, D. Bhattacharya, G. V. Shivashankar, *Biophys. J*. 2006, **91,** 2297-2303.

30. M. J. Hinner, K. Johnsson, *Curr. Opin. Biotech*.2010, **21,** 766–776.

31. B. Fierz, C.Chatterjee, R. K.McGinty, M. Bar-Dagan, D. P. Raleigh, T. W. Muir, *Nature Chem. Biol*. 2011, **7,** 113-119.

# Chapter 4

## Extending the fundamental imaging-depth limit of multi-photon microscopy by imaging with photo-activatable fluorophores

The content of this chapter has been published in:

Chen, Z.; Wei, L.; Zhu, X.; Min, W., “Extending the fundamental imaging-depth limit of multi-photon microscopy by imaging with photo-activatable fluorophores” *Opt. Express* **2012**, *20*, 18525. Z.C., L.W. and X. Z. contributed equally to this work.

## 4.1 Chapter outlook

It is highly desirable to be able to optically probe biological activities deep inside live organisms. By employing a spatially confined excitation via a nonlinear transition, multiphoton fluorescence microscopy has become indispensable for imaging scattering samples. However, as the incident laser power drops exponentially with imaging depth due to scattering loss, the out-of-focus fluorescence eventually overwhelms the in-focal signal. The resulting loss of imaging contrast defines a fundamental imaging-depth limit, which cannot be overcome by increasing excitation intensity. Herein we propose to significantly extend this depth limit by multiphoton activation and imaging (MPAI) of photo-activatable fluorophores. The imaging contrast is drastically improved due to the created disparity of bright-dark quantum states in space. We demonstrate this new principle by both analytical theory and experiments on tissue phantoms labeled with synthetic caged fluorescein dye or genetically encodable photoactivatable GFP.

I, Lu Wei, Xinxin Zhu and Wei Min are the main contributors of this work. Wei Min conceived the concept of MPAI. I and Lu Wei performed the spectroscopy and imaging experiments with photo-activatable fluorophores. Xinxin Zhu performed the experiments with photoactivatable GFP. Wei Min, I and Lu Wei wrote the paper with the contribution from Xinxin Zhu.

## 4.2 Introduction

Our ability to study the microscopic world has been revolutionized by advances in optical imaging technology. However, compared to electron microscopy and magnetic resonance imaging, optical microscopy suffers intrinsically from relatively coarse spatial resolution and superficial penetration depth. As one can imagine, resolution and penetration are two fundamentally coupled physical properties. While the diffraction-limited resolution barrier has been broken by a number of super-resolution fluorescence techniques such as STED, (F)PALM and STORM[1-4], the deepest penetration into scattering samples with sub-cellular resolution is achieved currently by multi-photon microscopy. By employing a nonlinear intensity-dependent optical excitation, multi-photon fluorescence is generated primarily at the focal volume where the probability of absorbing two (or more) photons by the same molecule at the same time is the highest, enabling an intrinsic 3D optical sectioning capability[5,6]. Such a spatially confined excitation scheme thus permits the capture of fluorescence photons emitted and then scattered from the focus by a wide-field non-descanned detector, thereby dramatically increasing the detection sensitivity of scattered signals. This profound feature ultimately leads to a significant imaging depth into scattering samples (more than three times deeper than that of one-photon confocal microscopy)[7,8].

As many other optical imaging modalities, a fundamental imaging-depth limit still exists for multiphoton fluorescence microscopy when imaging scattering samples labeled with fluorophores in 3D volume[9-13]. Here we take Figure 4-1 as an example: the two-photon fluorescence images of fluorescent beads embedded in a turbid 3D sample quickly vanish with the depth when using constant laser power (Figure 4-1a). This seemingly shallow cut-off depth is not the ultimate maximum, as imaging can actually be performed much deeper

provided that the incident laser power can be elevated accordingly to compensate for the scattering loss (Figure 4-1b). However, such a signal-promoting procedure comes at the cost of deteriorating imaging contrast: at a certain depth, the target beads can no longer be identified from the overwhelming background. Formally, the depth where the in-focus signal and the out-of-focus background are equal to each other is defined as the fundamental imaging-depth limit[10]. Note that the conventional optical sectioning picture that multiphoton fluorescence is generated only within the focal volume breaks down here, because the exponential increase of the incident laser power eventually outstrips the power-law fall-off of the excitation efficiency out-of-focus. Obviously, further increasing incident laser power cannot overcome this contrast-rendered imaging-depth limit.

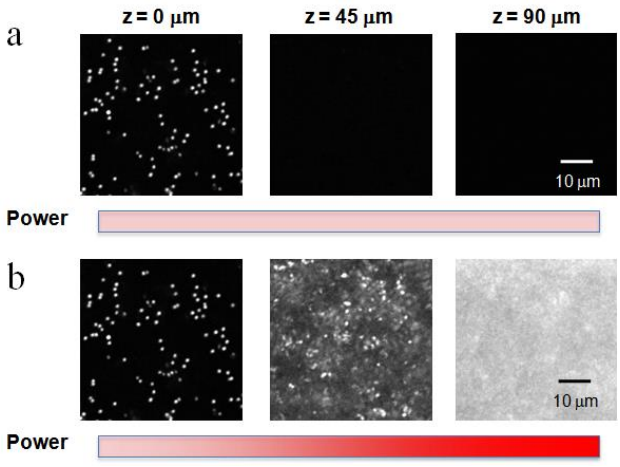



**Figure 4-1. Fundamental imaging-depth limit of multi-photon fluorescence microscopy.** (a) Images of a tissue phantom consisting of 5% intralipid, 1 μm fluorescent beads (Invitrogen F8765) and 1% agarose gel under a constant excitation laser power. Two-photon fluorescence signal quickly attenuates with the imaging depth. (b) Images of the same sample using a compensative higher laser power to maintain the signal strength at different depths. The resulting images, although showing signals deeper into the sample, suffer from a loss of contrast as the out-of-focus background grows. The fundamental imaging-depth limit is defined when the in-focus signal and the out-of-focus background are equal to each other.

Largely driven by the desire to perform *in vivo* deep tissue imaging, there have been tremendous efforts in improving the imaging depth of multiphoton microscopy. Several strategies have been explored, such as adaptive optics designed to pre-compensate for the scattering loss[14,15], imaging with longer excitation wavelength[16], chemical cleaning reagent[17], and differential aberration imaging[18]. However, relatively little work has been published on exploring novel imaging probes as a way of improving the image contrast or extending the imaging-depth limit.

Herein we present the application of photo-activatable fluorophores (PAFs) to extend the fundamental imaging-depth limit. PAFs are powerful imaging probes for tracking molecular and cellular dynamics with high temporal resolution, and have recently emerged as the key players in super-resolution fluorescence microscopy[2-4, 19-22] including nanoscopy with optical sectioning capability[23]. Unlike traditional fluorophores which always remain in their bright states, PAFs permit photo-induced transitions from dark states to bright states. By preferentially inducing the bright states into the focus and dark states into the out-of-focus background, multiphoton activation and imaging (MPAI) of PAFs can significantly enhance the imaging contrast and extend the depth limit of multiphoton microscopy. We demonstrate our proposal by both analytical theory and imaging experiments on tissue phantoms labeled with synthetic caged fluorescein dye or photoactivatable green fluorescent protein. Therefore, by offering the additional on-off molecular quantum states, PAFs seem to be capable of both breaking the diffraction-limited spatial resolution and extending the imaging-depth limit.

## 4.3 Results and discussions

### 4.3.1 The fundamental imaging-depth limit of two-photon microscopy

Let us first frame the problem of deep two-photon imaging with regular fluorophores. The fundamental imaging-depth limit can be defined as the point where the in-focus signal and the out-of-focus background are equal[9-13]:

$$\left(\frac{S}{B}\right)_{regular} = \frac{\int_{V_{in}}\int_0^{\tau} C_S(r,z) I_i^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V}{\int_{V_{out}}\int_0^{\tau} C_B(r,z) I_i^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V} = 1, \qquad (4\text{-}1)$$

where $V_{in}$, is the focal volume at the focal plane, $V_{out}$, is the total sample volume along the light path but excluding the volume at the focal plane, $r$, is the distance from the optical axis, $z$, is the axial distance from the tissue surface, $C$, is the local fluorophore concentration, $I_i$ is the local imaging laser intensity, and $\tau$, is the pixel dwell time during imaging. We assume that there is no fluorophore saturation or photobleaching and that the fluorescence collection efficiencies at the wide-field detector are identical for the signal and the background.

We can now analyze the laser intensity distribution within a scattering sample. In a typical scenario where the fluorophores are distributed throughout the sample volume, the number of out-of-focus fluorophores is almost always much larger than that of the in-focus ones: $\int_{V_{out}} C_B(r,z)\,\mathrm{d}V >> \int_{V_{in}} C_S(r,z)\,\mathrm{d}V$. Therefore, at the fundamental imaging-depth limit as defined in Eq. (4-1), $I_i^2(r,z)\big|_{V_{in}}$ at the focus should be much larger than the out-of-focus counterpart $I_i^2(r,z)\big|_{V_{out}}$, despite of the scattering loss. Indeed, in the simple condition of homogeneous fluorophore distribution, i.e., $C_B(r,z) = C_S(r,z)$, $\int_V\int_0^{\tau} I_i^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V$ will be identical between the background and the signal, and consequently, the integral of $I_i^2(r,z)$

over a subset layers of the out-of-focus volume (e.g., $\int_{V_{out}} I_i^2(r,z)\,\mathrm{d}\,r$ for any $z$) will be smaller than that over the focus.

### 4.3.2 Reducing the background fluorophore concentration alone can improve the depth limit

Although the sample scattering is seemingly the origin of the limited light penetration, the above theoretical framework suggests that the concentration of background fluorophores should play an important role in determining the depth limit. To separate these two effects, we constructed and imaged a set of "two-layer" samples (Figure 4-2). 0-5 % intralipid emulsions were used as they closely mimic the response of biological tissues to near infrared light. In the absence of scattering intralipid, increasing fluorescein concentration in the background does not affect the two-photon imaging quality of the target fluorescent beads. In contrast, when the scattering effects are strong enough (e.g., 2% and 5% intralipid), the two-photon imaging contrast becomes anti-correlated with the background fluorescein concentration. Thus, both strong sample scattering and abundant background fluorophores are necessary conditions for the limited imaging-depth. Sample scattering alone does not create poor imaging contrast. As an important insight, images presented in Figure 4-2 suggest that, simply reducing the background dye concentration while maintaining the high intralipid content still allows imaging the target beads with increased contrast.

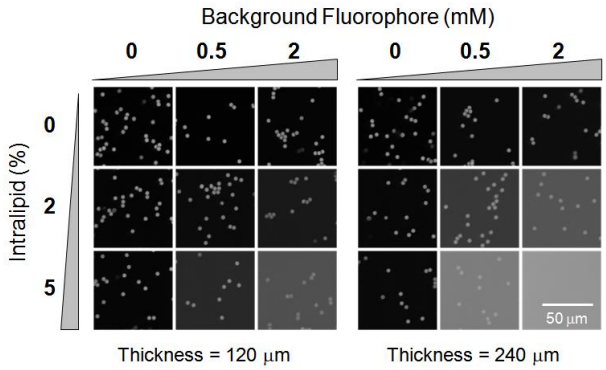



**Figure 4-2. Imaging contrast of multi-photon microscopy depends on both sample scattering and background fluorophore concentration.** A set of "two-layer" samples (fluorescent beads were placed on a glass coverslip as the target, while a thick layer of mixed fluorescein dye solution and scattering intralipids was inserted as the background between the target and the objective.) with varying intralipid contents and background dye concentrations were imaged using a two-photon microscope. The laser power was set higher accordingly when a more scattering sample was imaged. At a given background layer thickness, image contrast deteriorates only when both a significant background turbidity and a dense background fluorophore staining are present. Imaging contrast further deteriorates when thicker background layers were applied.

### 4.3.3 Theoretical framework of MPAI using dynamic PAFs

Inspired by the above insight, we envision a way to preferentially keep the background fluorophores in the dark states by using PAFs which have to be activated first by multiphoton process before excited by another multiphoton process. Two different modes of experiments are possible: simultaneous or sequential activation and imaging. When PAFs are being activated and imaged by $I_a$ and $I_i$ (one same activation-imaging laser or two combined lasers, where $I_a$ is the local intensity of the activation laser) simultaneously, $C(r,z)$ in Eq. (4-1) should be replacecd by $A(r,z,t)$, the time-dependent concentration of the PAFs in the bright state. Quantitatively, $A(r,z,t)$ will be the product of $C(r,z)$ and the time-dependent multiphoton activation yield, $\eta(t)$:

$$A(r,z,t) = C(r,z)\eta(r,z,t). \tag{4-2}$$

In the simplest condition, $\eta(t)$ follows a first-order chemical kinetics:

$$\eta(t) = 1 - \exp\left[-\sigma I_a^2(r,z,t)t\right] \quad 0 \le t \le \tau, \tag{4-3}$$

with a rate constant being proportional to $I_a^2(r,z)$ of the activation laser and the two-photon activation cross section, $\sigma$. Alternatively, we describe the sequential mode in which activation by $I_a$ and imaging by $I_i$ are performed at two laser wavelengths in sequential raster scans: $A(r,z) = C(r,z)\eta(r,z,\psi)$, where $\psi$ is the pixel dwell time of the multiphoton activation scanning process, and the resulting $A(r,z)$ becomes time independent to the subsequent imaging process.

When the timescale for PAFs to diffuse or transport out of the laser focal volume is shorter than the frame acquisition time, the accumulation of the activated fluorophores in the background during the scanning process could be neglected. This dynamic condition is rather common, as exemplified by *in vivo* blood vessel imaging in which fluorophores are injected into the circulatory systems of small animals[24]. Hence, the signal-to-background ratio for dynamic PAFs can be expressed as

$$\left(\frac{S}{B}\right)_{PAFs} = \frac{\int_{V_{in}}\int_0^{\tau} C_S(r,z)\left[1-e^{-\sigma I_a^2(r,z,t)t}\right] I_i^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V}{\int_{V_{out}}\int_0^{\tau} C_B(r,z)\left[1-e^{-\sigma I_a^2(r,z,t)t}\right] I_i^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V}. \tag{4-4}$$

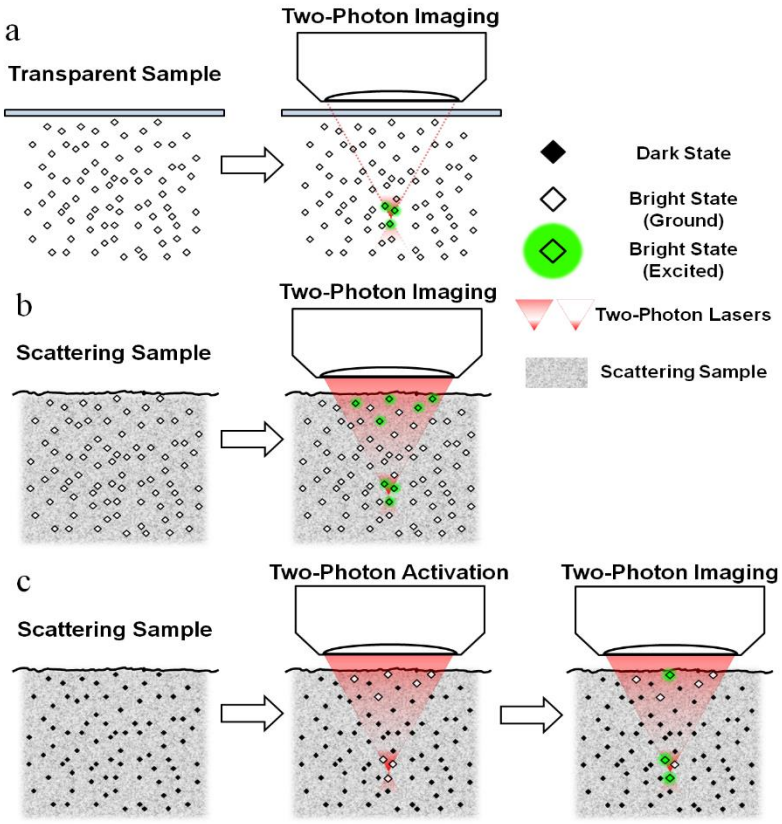



**Figure 4-3. Principle of multiphoton activation and imaging (MPAI).** (a) When imaging transparent samples, fluorescence is only generated at the laser focus where the intensity is the highest. (b) When imaging deep into scattering samples, substantial laser intensities are distributed out of focus, generating background fluorescence that is comparable to or even stronger than the in-focus signal. (c) When imaging with PAFs which are originally in the dark state, the multiphoton activation will switch on a higher percentage of PAFs at focus than those out-of-focus. Such a spatial disparity of dark-bright transitions would lead to a significantly decreased background fluorescence in the subsequent imaging.

As was analyzed earlier, the laser intensity at the focus is much higher than its out-of-focus counterparts at the imaging-depth limit defined by Eq. (4-1). This would lead to a much higher dark-to-bright conversion yield at the focus: $\eta_S(r,z,t) > \eta_B(r,z,t)$, under a proper pixel dwell time with a non-saturating activation yield. Consequently, we expect a much improved signal-to-background ratio for PAFs than for regular fluorophores:

$$(S/B)_{PAFs} > (S/B)_{regular}. \qquad (4\text{-}5)$$

It is constructive to note that, after the photoactivation (with a non-saturating activation yield), the number of in-focus activated PAFs become comparable to the total number of the activated ones out-of-focus:

$$\left(\frac{N_S}{N_B}\right)_{activated} = \frac{\int_{V_{in}}\int_0^{\tau} C_S(r,z)\left[1-e^{-\sigma I_a^2(r,z,t)\tau}\right]\mathrm{d}t\,\mathrm{d}V}{\int_{V_{out}}\int_0^{\tau} C_B(r,z)\left[1-e^{-\sigma I_a^2(r,z,t)\tau}\right]\mathrm{d}t\,\mathrm{d}V} \approx \frac{\int_{V_{in}}\int_0^{\tau} C_S(r,z) I_a^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V}{\int_{V_{out}}\int_0^{\tau} C_B(r,z) I_a^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V} = 1. \quad (4\text{-}6)$$

Compared to the regular dye case in which:

$$\left(\frac{N_S}{N_B}\right)_{regular} = \frac{\int_{V_{in}} C_S(r,z)\,\mathrm{d}V}{\int_{V_{out}} C_B(r,z)\,\mathrm{d}V} \approx \frac{V_{in}}{V_{out}} << 1. \quad (4\text{-}7)$$

Eq. (4-6) clearly indicates that the bright states are preferentially induced at the focus whereas the out-of-focus PAFs mostly remain in dark states. As shown in Figure 4-3, such a remarkable disparity of bright-dark states in space would significantly reduce the background contribution from out-of-focus fluorophores.

### 4.3.4 Experimental demonstration of MPAI using a caged organic dye

Caged fluorescein, resorufin and rhodamine are the first developed small organic PAFs. They have been applied to study the assembly of tubulin, hydrodynamic flows and cell lineage during embryo development[25]. Here we use the caged fluorescein, which is commercially available, as a proof-of-principle. Figure 4-4(a) shows the photo-uncaging reaction of caged fluorescein. The dark state can be activated to the bright state by illuminating with UV light as shown by the absorption and fluorescence spectra and uncaging kinetics in Figure 4-4(b). For our two-photon application, laser pulses at 750 nm can uncage the protective groups and subsequently excite the bright state. Therefore only one laser is needed in this simultaneous mode, i.e., $I_a = I_i$. We again constructed "two-layer"

samples. The target object, a droplet of caged fluorescein solution, was placed on a glass coverslip, and a thick layer made of mixed scattering beads and caged fluorescein was inserted between the target and the objective. Control two-layer samples exhibit same physical parameters except for the use of regular fluorescein in both the target droplets and the out-of-focus background layers.

Figure 4-4(c)-(f) show the comparison between $(S/B)_{regular}$ and $(S/B)_{PAFs}$ in our tissue phantoms labeled with dynamic fluorophores. Using control samples, we first determined the proper concentrations of scattering and dye species in the background layer so that imaging with regular fluorescein is approaching the fundamental imaging-depth limit with $(S/B)_{regular} \approx 1.2$ (Figure 4-4c). We then switched to the corresponding PAFs sample. To accommodate the relatively slow photo-activation of caged fluorescein, a long pixel dwell time of ~ 1.0 ms was used to obtain enough fluorescence signals. The resulting $(S/B)_{PAFs}$ is found to be about 20 in Figure 4-4(d), exhibiting a significant improvement of imaging contrast. Encouraged by this, we further elevated the background dye concentration by a factor of 3 for both control and PAFs samples in Figure 4-4(e) and 4-4(f). As expected, $(S/B)_{regular} \approx 1/3$ is now beyond what the depth limit permits. In contrast, the corresponding PAFs sample exhibits a remarkable $(S/B)_{PAFs} \sim 8$.

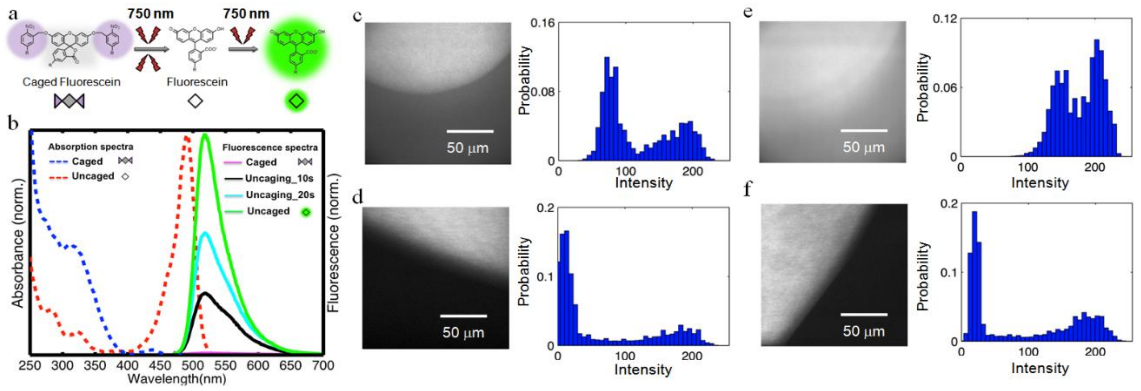



**Figure 4-4. Experimental demonstration of MPAI with caged fluorescein.** (a) CMNB (5-carboxymethoxy-2-nitrobenzyl) caging groups of CMNB-caged fluorescein could be cleaved by a 750 nm pulsed laser, leading to the formation of fluorescein in the bright state, which could be excited by the same laser to emit fluorescence. (b) Absorption and fluorescence spectra of caged fluorescein and uncaged fluorescein, with the activation kinetics under UV illumination. A ~100-fold fluorescence enhancement was observed after a complete activation. (c-f) Imaging "two-layer" samples (Fig. 7.) where the targets are 2 μL droplets of 1 mM dye solution and the background layers (120 μm thick) consist of scattering polystyrene beads (0.9 μm) and dye solution (1 mM for (c) and (d), 3 mM for (e) and (f)). Boundaries of liquid droplets cross the field of view so that the darker parts correspond to the background while the brighter parts represent the sum of the signal and the background. When the imaging depth-limit is reached for regular fluorescein in (c) with a S/B of about 1.2, a 20-times improvement is achieved for caged fluorescein in (d). With a three folds more dyes in the background layer of (e) and (f), imaging contrast becomes extremely poor for regular fluorescein, while the caged fluorescein still offers a S/B of about 8.

### 4.3.5 Theoretical framework of MPAI using static PAFs

If PAFs are rather static during the entire frame acquisition, we then need to consider the spatio-temporal accumulation of the activated fluorophore in out-of-focus background during the raster scanning. In this scenario, while $\eta_S$ of the signal remains the same as in Eqs. (4-2)-(4-4), the effective activation time for $\eta_B$ will be longer than the pixel dwell time, $\tau$, by a factor of, $f(z)$, which scales with the $z$-dependent laser beam area. The result of static PAFs hence becomes:

$$\left(\frac{S}{B}\right)_{PAFs} = \frac{\int_{V_{in}}\int_0^{\tau} C_S(r,z)\eta_S(r,z,t)I_i^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V}{\int_{V_{out}}\int_0^{\tau} C_B(r,z)\eta_B\left(r,z,f(z)\cdot\tau\right)I_i^2(r,z,t)\,\mathrm{d}t\,\mathrm{d}V}. \quad (4\text{-}8)$$

Thus, due to the accumulation of the activated fluorophores during laser scanning, the out-of-focus background will be higher here compared to that from dynamic PAFs described in Eq. (4-4). In the condition of low photo-activation yield, we have the following approximation

$$\frac{\eta_S(r,z,\tau)}{\eta_B\left(r,z,f(z)\cdot\tau\right)} \approx \frac{I_a^2(r,z)\big|_{V_{in}}}{f(z)\cdot I_a^2(r,z)\big|_{V_{out}}} \approx \frac{\int_{in} I_a^2(r,z)\,\mathrm{d}r}{\int_{out} I_a^2(r,z)\,\mathrm{d}r} > 1. \quad (4\text{-}9)$$

From the earlier analysis of laser intensity distribution at the depth limit defined in Eq. (4-1), we have inferred that $\int_{in} I^2(r,z)\,\mathrm{d}r > \int_{out} I^2(r,z)\,\mathrm{d}r$ for all $z$ layers. Consequently, the final activation yield will be higher at the focus than in the background. Therefore, the signal-to-background ratio using static PAFs is still higher in Eq. (4-8) than that with regular fluorophores in Eq. (4-1).

### 4.3.6 Experimental demonstration of MPAI using a photo-activatable fluorescent protein

PAFs also include several fluorescent proteins that can be genetically encoded and fused to other proteins of interest for *in vivo* imaging of cells and animals. Photoactivatable green fluorescent protein (pa-GFP) is the first member of this family of optical highlighters[26]. As shown in Figure 4-5(a), pa-GFP can be turned on from its initial dark state by two-photon activation in the 750~850 nm range, and the resulting bright state can be readily imaged in the 900~950 nm range[27].

We prepared tissue phantoms by embedding the *E. coli* cells expressing fluorescent protein into 3D agarose gel. The resulting sample is highly scattering due to the densely packed *E. coli* cells. The PAFs in this sample are considered static, as the cells are stationary within the agarose gel. As demonstrated in Figure 4-5(b) and 4-5(c), while the out-of-focus background is overwhelming for cells expressing regular GFP at a depth of 100 μm, MPAI of cells expressing pa-GFP at the same depth offers a satisfactory image contrast. Thus, we have experimentally demonstrated the ability of MAPI in extending the fundamental imaging-depth limit of two-photon fluorescence microscopy.

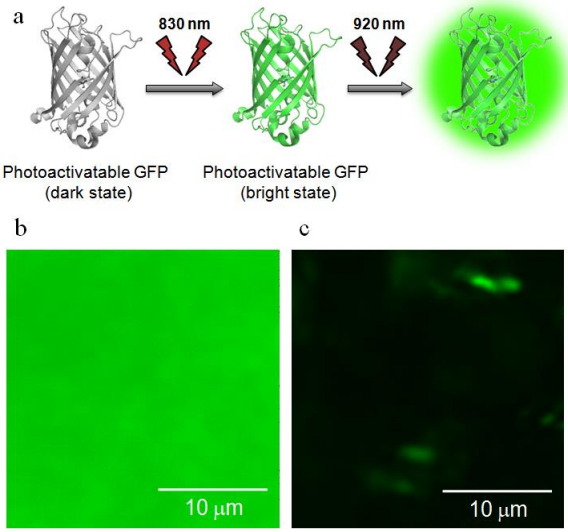



**Figure 4-5. Experimental demonstration of MPAI with pa-GFP**. (a) pa-GFP could be activated by a pulsed laser at 830 nm to its bright state, which could be further excited by a 920 nm pulsed laser to emit fluorescence. (b, c) Deep imaging comparison of 3D turbid samples made of *E. coli* cells expressing free regular GFP (b) or pa-GFP (c) embedded in 2% agarose gel with the same cell densities. While out-of-focus background is overwhelming when imaging *E. coli* expressing regular GFP at a 100 μm depth inside the gel, MPAI with pa-GFP at the same depth offers a satisfactory image contrast.

## 4.4 Methods and materials

### 4.4.1 Sample preparation

Fluorescent beads were purchased from Invitrogen (F8765: green/yellow 1.0 μm beads; F8859: green/yellow 4.0 μm beads). Fluorescein (Sigma 46960) and CMNB-Caged Carboxyfluorescein (Invitrogen C-20050) were prepared into 10 mM stock solution in DMSO before diluting to final concentrations. Intralipid (Sigma I141) or non-fluorescent polystyrene beads (Sigma CLB9) were used as the scattering species. "Two-layer" samples in Figures 4-2 and 4-4 were prepared as shown in Figures 4-6 and 4-7: An imaging spacer with a 120 μm thickness (Sigma S7935) was pasted onto a piece of micro coverglass (VWR 48393-220). Inside the spacer, 5 μL of material serving as the scattering background was added. For the activation-imaging experiments of fluorescein and caged fluorescein (Figure 4-4c-f), suspension of non-fluorescent 0.9 μm polystyrene beads (Sigma CLB9) was sonicated and diluted (1:5) into 50% glycerol solution with fluorescein or caged fluorescein at a concentration of 1 mM or 3 mM to serve as the scattering background. The background material was then sealed inside the spacer by another piece of micro coverglass on top, giving a sandwich-like background sample. Imaging targets (fluorescent beads or dye droplets) were placed on top of the glass sandwich before imaging. Glycerol solutions were used to minimize evaporation of solvent during the imaging process. In bacteria imaging experiments, BL21 *E. coli* strains expressing pa-GFP (Addgene plasmid 11911) and Dronpa-3 (a regular GFP version for control imaging, MBL International Corporation) were harvested by centrifugation, resuspended and embedded inside an imaging spacer using 2% agarose gel. *E. coli* densities of different samples were calibrated based on OD600 measurement to achieve the consistent scattering effect.

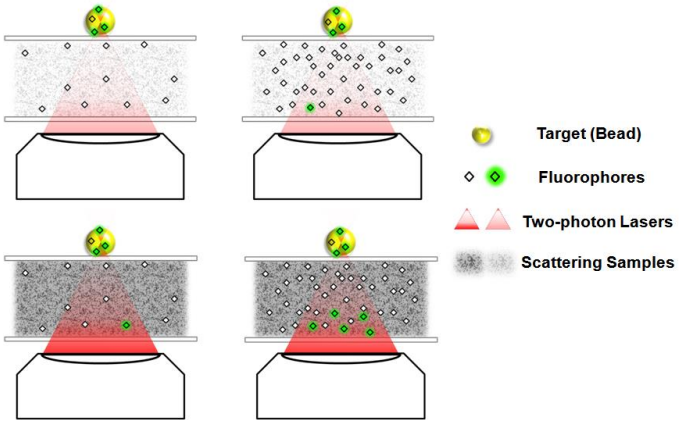



**Figure 4-6: The "two-layer" sample design for beads imaging experiments**. Fluorescent beads were placed on a glass coverslip as the target, while a thick layer of mixed fluorescein dye solution with scattering intralipid was inserted as the background between the target and the objective. From left to right, the background fluorescein concentration increases; from top to bottom, the scattering intralipid percentage increases. The signal-to-background ratio decreases with the increase of both the background scattering and the background dye concentration.

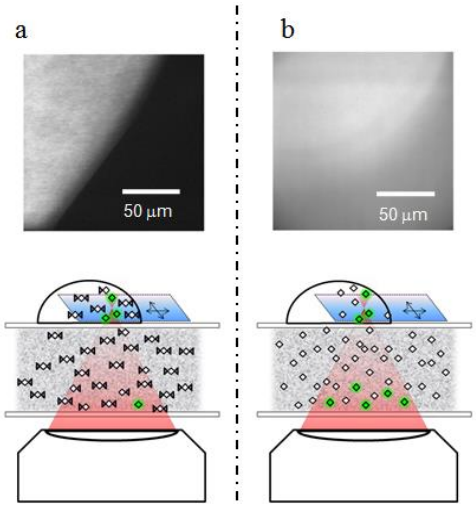



**Figure 4-7: The "two-layer" sample design for droplet imaging experiments**. (a): imaging with caged fluorescein (Figure 4-4f). (b): imaging with regular fluorescein (Figure 4-4e). Images were taken across the boundaries of caged-fluorescein or regular fluorescein droplets on top of a layer of caged-fluorescein or regular fluorescein solution doped with scattering polystyrene beads.

### 4.4.2 Fluorescence imaging

All the two-photon fluorescence images were taken on a Leica TCS SP5 MP inverted microscope equipped with a Mai Tai HP laser (690-1040 nm tunable emission) and a HCX

PL APO CS 20X DRY microscope objective (N.A. 0.70). A non-descanned epi PMT detector placed directly behind the objective was used for fluorescence collection in combination with a 680/SP emission filter. Images were processed with LAS AF software unless otherwise stated. Images of fluorescent beads (Figures 4-1 and 4-2) were taken under 750 nm with a pixel dwell time of ~2.4 μs. Images of caged-fluorescein-based samples (Figure 4-4d and 4-4f) were taken under 750 nm (48mW) with a pixel dwell time of 977 μs to achieve effective uncaging. Images of *E. coli* expressing regular GFP (Figure 4-5b) were taken under 920 nm (65mW) with a pixel dwell time of 98 μs. Images of *E. coli* expressing pa-GFP (Fig. 5c) were taken under 920 nm (65mW) after 3 frames of two-photon activation under 830 nm (95 mW) with a pixel dwell time of 98 μs. Note that for tissue phantoms labeled with caged fluorescein or pa-GFP, due to the sample scattering, the actual laser power reached at the deep focal plane is much lower than the measured total power. All images were acquired with 512 by 512 pixels.

## 4.5 Conclusion

To summarize, we have presented the application of PAFs in extending the fundamental imaging-depth limit of multiphoton fluorescence microscopy. This represents a novel strategy compared to the existing approaches that focus on reducing sample scattering. Theoretical framework has been provided to describe both the dynamic and static PAFs during image acquisition. Experimentally, we demonstrated MPAI on scattering tissue phantoms labeled with caged fluorescein (dynamic PAFs scenario) or pa-GFP (static PAFs scenario), which serves as an example of synthetic small-molecule and genetically encoded PAF, respectively.

The practical performance of MPAI can be potentially improved by using PAFs with better designed photophysical properties. First, two-photon photo-activation could be much faster for molecules with larger activation cross sections. For instance, caged coumarin has been shown to exhibit 100 times faster uncaging rate via substrate-assisted photolysis[28]. Second, after being photo-activated, many photochromic fluorophores can also be switched off either thermally or by light[23, 29-30]. Such a switching-off property could re-set PAFs in the entire sample and hence allow repeated imaging on the same x-y plane or navigating through multiple layers.

It does not seem to be a mere coincidence that PAFs can play key roles for both breaking the diffraction-limited spatial resolution of fluorescence microscopy and extending the fundamental imaging-depth limit of multiphoton microscopy. In parallel, it is noteworthy to compare this work to stimulated emission reduced fluorescence (SERF) microscopy which was recently proposed for extending the fundamental imaging-depth limit of two-photon imaging[31]. While SERF microscopy is reminiscent of STED microscopy, the PAFs approach presented here can be considered as being related to PALM in terms of harnessing photo-activation. In essence, the coarse spatial resolution and the superficial penetration depth are both rooted in the difficulty of distinguishing identical molecules in space. By offering the additional disparity of (on-off) molecular quantum states, PAFs render the capability of circumventing both resolution and penetration limits.

## 4.6 References

1. S. W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimulated emission: stimulated emission-depletion fluorescence microscopy,” Opt. Lett. **19**, 780–782 (1994).

2. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution, " Science **313**, 1642-1645 (2006).
3. S. T. Hess, T. P. K. Giririjan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," Biophys. J. **91**, 4258-4272 (2006).
4. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)", Nature Methods **3**, 793-795 (2006).
5. W. Denk, J. H. Strickler, and W. W. Webb, "Two-photon laser scanning microscopy," Science **352**, 73-76 (1990).
6. W. R. Zipfel, R. M. Williams, and W. W. Webb, "Nonlinear magic: multiphoton microscopy in the biosciences," Nat. Biotechnol. **21**, 1369-1377 (2003).
7. F. Helmchen and W. Denk, "Deep tissue two-photon microscopy," Nat. Methods **2**, 932-940 (2005).
8. R. Yuste. Ed. *Imaging: a laboratory manual* (Cold Spring Harbor Laboratory Press, 2010)
9. P. Theer, M. T. Hasan, and W. Denk, "Two-photon imaging to a depth of 1000 microm in living brains by use of a Ti:Al2O3 regenerative amplifier," Opt. Lett. **28**, 1022-1024 (2003).
10. P. Theer and W. Denk, "On the fundamental imaging-depth limit in two-photon microscopy," J. Opt. Soc. Am. A-Opt. Image Sci. Vis. **23**, 3139-3149 (2006).
11. D. Kobat, N. G. Horton, and C. Xu, "In vivo two-photon microscopy to 1.6-mm depth in mouse cortex," J. Biomed. Opt. **16**, 106014 (2011).
12. N. J. Durr, C. T. Weisspfennig, B. a Holfeld, and A. Ben-Yakar, "Maximum imaging depth of two-photon autofluorescence microscopy in epithelial tissues," J. Biomed. Opt. **16**, 026008 (2011).
13. J. Ying, F. Liu, and R. R. Alfano, "Spatial distribution of two-photon-excited fluorescence in scattering media," Appl. Opt. **38**, 224-229 (1999).
14. N. Ji, D. E. Milkie, and E. Betzig, "Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues," Nat. Methods **7**, 141-147 (2010).
15. M. Rueckel, J. a Mack-Bucher, and W. Denk, "Adaptive wavefront correction in two-photon microscopy using coherence-gated wavefront sensing," Proc. Natl. Acad. Sci. USA **103**, 17137-17142 (2006).
16. D. Kobat, M. E. Durst, N. Nishimura, A. W. Wong, C. B. Schaffer, and C. Xu, "Deep tissue multiphoton microscopy using longer wavelength excitation," Opt. Express **17**, 13354-13364 (2009).
17. H. Hama, H. Kurokawa, H. Kawano, R. Ando, T. Shimogori, H. Noda, K. Fukami, A. Sakaue-Sawano, and A. Miyawaki, "Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain," Nat. Neurosci. **14**, 1481-1488 (2011).

18. A. Leray, K. Lillis, and J. Mertz, "Enhanced background rejection in thick tissue with differential-aberration two-photon microscopy," Biophys. J. **94**, 1449-1458 (2008).

19. M. Fernández-Suárez and A. Y. Ting, "Fluorescent probes for super-resolution imaging in living cells," Nat. Rev. Mol. Cell Biol. **9**, 929-943 (2008).

20. J. Lippincott-Schwartz and G. H. Patterson, "Photoactivatable fluorescent proteins for diffraction-limited and super-resolution imaging," Trends Cell Biol. **19**, 555-565 (2009).

21. M. Heilemann, P. Dedecker, J. Hofkens, and M. Sauer, "Photoswitches: key molecules for subdiffraction-resolution fluorescence imaging and molecular quantification," Laser Photon. Rev. **3**, 180-202 (2009).

22. H.-l. D. Lee, S. J. Lord, S. Iwanaga, K. Zhan, H. Xie, J. C. Williams, H. Wang, G. R. Bowman, E. D. Goley, L. Shapiro, R. J. Twieg, J. Rao, and W. E. Moerner, "Superresolution imaging of targeted proteins in fixed and living cells using photoactivatable organic fluorophores," J. Am. Chem. Soc. **132**, 15099-15101 (2010).

23. J. Fölling, V. Belov, R. Kunetsky, R. Medda, A. Schönle, A. Egner, C. Eggeling, M. Bossi, and S. W. Hell, "Photochromic rhodamines provide nanoscopy with optical sectioning," Angew. Chem. Int. Ed. **46**, 6266-6270 (2007).

24. D. Kleinfeld, P. P. Mitra, F. Helmchen, and W. Denk, "Fluctuations and stimulus-induced changes in blood flow observed in individual capillaries in layers 2 through 4 of rat neocortex," Proc. Natl. Acad. Sci. USA **95**, 15741-15746 (1998).

25. T. Mitchison, K. Sawin, J. Theriot, and K. Gee, "Caged fluorescent probes," Methods Enzymol. **291**, 63-78 (1998).

26. G. H. Patterson and J. Lippincott-Schwartz, "A photoactivatable GFP for selective photolabeling of proteins and cells," Science **297**, 1873-1877 (2002).

27. M. Schneider, S. Barozzi, I. Testa, M. Faretta, and A. Diaspro, "Two-photon activation and excitation properties of PA-GFP in the 720-920-nm region," Biophys. J. **89**, 1346-1352 (2005).

28. Y. Zhao, Q. Zheng, K. Dakin, K. Xu, M. L. Martinez, and W.-H. Li, "New caged coumarin fluorophores with extraordinary uncaging cross sections suitable for biological imaging applications," J. Am. Chem. Soc. **126**, 4653-4663 (2004).

29. R. Ando, H. Mizuno, and A. Miyawaki, "Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting," Science **306**, 1370-1373 (2004).

30. M.-Q. Zhu, G.-F. Zhang, C. Li, M. P. Aldred, E. Chang, R. a Drezek, and A. D. Q. Li, "Reversible two-photon photoswitching and two-photon imaging of immunofunctionalized nanoparticles Targeted to Cancer Cells," J. Am. Chem. Soc. **133**, 365-372 (2010).

31. L. Wei, Z. Chen, and W. Min, "Stimulated emission reduced fluorescence microscopy: a concept for extending the fundamental depth limit of two-photon fluorescence imaging," Biomed. Opt. Express **3**, 1465-1475 (2012).

# Chapter 5

# Multicolor Live-Cell Chemical Imaging by Isotopically Edited Alkyne Vibrational Palette

The content of this chapter has been published in:

Chen, Z.; Paley, D. W.; Wei, L.; Weisman, A. L.; Friesner, R. A.; Nuckolls, C.; Min, W., “Multicolor Live-Cell Chemical Imaging by Isotopically Edited Alkyne Vibrational Palette” *J. Am. Chem. Soc.* **2014**, *136*, 8027

## 5.1 Chapter outlook

Vibrational imaging such as Raman microscopy is a powerful technique for visualizing a variety of molecules in live cells and tissues with chemical contrast. Going beyond the conventional label-free modality, recent advance of coupling alkyne vibrational tags with stimulated Raman scattering microscopy paves the way for imaging a wide spectrum of alkyne-labeled small biomolecules with superb sensitivity, specificity, resolution, biocompatibility, and minimal perturbation. Unfortunately, the currently available alkyne tag only processes a single vibrational "color", which prohibits multiplex chemical imaging of small molecules in a way that is being routinely practiced in fluorescence microscopy. Herein we develop a three-color vibrational palette of alkyne tags using a $^{13}C$-based isotopic editing strategy. We first synthesized $^{13}C$ isotopologues of EdU, a DNA metabolic reporter, by using the newly developed alkyne cross-metathesis reaction. Consistent with theoretical predictions, the mono-$^{13}C$ ($^{13}C{\equiv}^{12}C$) and bis-$^{13}C$ ($^{13}C{\equiv}^{13}C$) labeled alkyne isotopologues display Raman peaks that are red-shifted and spectrally resolved from the originally unlabeled ($^{12}C{\equiv}^{12}C$) alkynyl probe. We further demonstrated three-color chemical imaging of nascent DNA, RNA, and newly uptaken fatty-acid in live mammalian cells with a simultaneous treatment of three different isotopically edited alkynyl metabolic reporters. The alkyne vibrational palette presented here thus opens up multicolor imaging of small biomolecules, enlightening a new dimension of chemical imaging.

I have been the main contributor of the project. I conceived the project. Andrew Weisman performed the DFT calculations under the supervision of Richard Friesner. Daniel Paley and I developed and optimized the metathesis chemistry. Lu Wei built the SRS microscope and developed live cell imaging protocols. I performed other parts of chemical

synthesis, cell spectroscopy/imaging experiments and data analysis. Colin Nuckolls supervised the metathesis chemistry and Wei Min supervised the overall project. Wei Min and I wrote the paper with contributions from all co-authors.

## 5.2 Introduction

Advances in optical microscopy in the past decades have revolutionized the way modern biological sciences are conducted. In particular, powerful fluorescence imaging techniques have flourished, largely driven by the advent of a diversity of fluorescent probes including organic dyes[1], genetically encoded fluorescent proteins[2] and semiconductor quantum dots[3]. A recurring theme in all these exciting developments is the creation of a palette of multiple colors resolvable from each other in the visible spectrum. Through targeting several species of interest simultaneously, these palette sets have enabled multiplex studies for visualization, localization and interaction in a broad spectrum of structural and functional assays. Notable applications include protein-protein interactions by FRET between different fluorescent proteins or organic dyes[4], super-resolution structural imaging by multi-color STED[5], PALM/STORM[6], as well as functional imaging using palettes of calcium-sensitive proteins[7] and voltage-sensitive dyes[8]. However, fluorescent probes are not suitable for tagging small biomolecules (e.g., nucleic acids, amino acids, fatty acids), because the relatively bulky fluorescent tags (even the smallest dyes) often destroy or significantly alter the biological activities of small biomolecules.

Raman-based vibrational microscopy represents an alternative to fluorescence microscopy. Raman microscopy is well suited for probing small biomolecules, especially when coupled with specific, small-size vibrational tags. Among the existing vibrational tags such as bioorthogonal chemical moieties or stable isotopes[9], alkynes are unique due to a combination of several merits including its small size (two atoms), high Raman activity of C≡C stretching and a signal frequency well separated from endogenous cellular background. Bioimaging using alkyne tags was initially demonstrated with 5-ethylnyldeoxyuridine (EdU),

a cell-proliferation reporter, and other mobile molecules by conventional spontaneous Raman microscopy [10,11]. Very recently, our group have coupled alkyne tags with stimulated Raman scattering (SRS) microscopy as a general strategy for imaging small biomolecules[12] (Figure 5-1a). Compared to the spontaneous counterpart, SRS is a state-of-the-art chemical imaging technique offering: substantial signal amplification, sensitivity increases, speed acceleration, immunity to auto-fluorescence, and optical penetration and sectioning in 3D tissues and whole animals[13–16]. As such, we have achieved visualizing metabolic incorporation of alkyne-tagged small precursors of deoxyribonucleoside, ribonucleoside, amino acid, choline and fatty acid into newly synthesized DNA, RNA, proteomes, phospholipids and triglycerides, respectively, in live cells and organisms, and tracking 3D delivery of an alkyne-bearing drug in mouse skin tissue[12].

In terms of specific labeling and detection, Raman imaging of alkyne-tagged small biomolecules is conceptually analogous to fluorescence imaging of fluorophore (including dyes, proteins and quantum dots) labeled larger species. However, unlike its fluorescence counterpart, multi-color Raman imaging of alkyne-tagged molecules lacks a general solution. The Raman vibrational frequencies of alkynyl molecules, located in a spectral region between 2080 $cm^{-1}$ to 2260 $cm^{-1}$, depend on the chemical structures and are not easily subject to customization. Examples of dual-color Raman imaging of alkyne-tagged molecules have been demonstrated using two structurally different alkynyl probes in which the electronic or conjugation properties of the triple bonds render resolvable Raman peaks[11,12]. However, Raman peaks of most of typical alkynyl probes still overlap with each other. Therefore, the prefixed Raman frequencies of alkyne-tagged molecules prevents multiplexed chemical

imaging of small molecules in a general way that is being routinely practiced in fluorescence microscopy.

The study reported here reveals a general chemical strategy to expand the vibrational palette of terminal alkyne tags. Inspired by classical isotope approaches adapted for vibrational spectroscopy[17–22], we envision that introduction of one or two heavy $^{13}C$ atom(s) into the alkynyl group would dampen the original stretching frequency of the $^{12}C{\equiv}^{12}C$ bond and thus create new vibrational "colors". To address the underlying synthetic challenge, we show here the preparation of three distinct forms of $^{13}C$ isotopically edited alkyne vibrational tags by using the newly developed alkyne cross-metathesis chemistry[23]. The three forms of alkynes ($^{12}C{\equiv}^{12}C$; $^{13}C{\equiv}^{12}C$; $^{13}C{\equiv}^{13}C$) are biochemically identical, and, to our delight, display three mutually resolvable Raman peaks. We demonstrate three-color SRS imaging of DNA, RNA and lipid metabolism using three different alkyne-tagged small-molecule metabolic reporters in live mammalian cells. Thus, our isotopic editing approach for creating alkyne palettes paves the way for multicolor chemical imaging, bringing small biomolecules under the illumination of modern light microscopy.

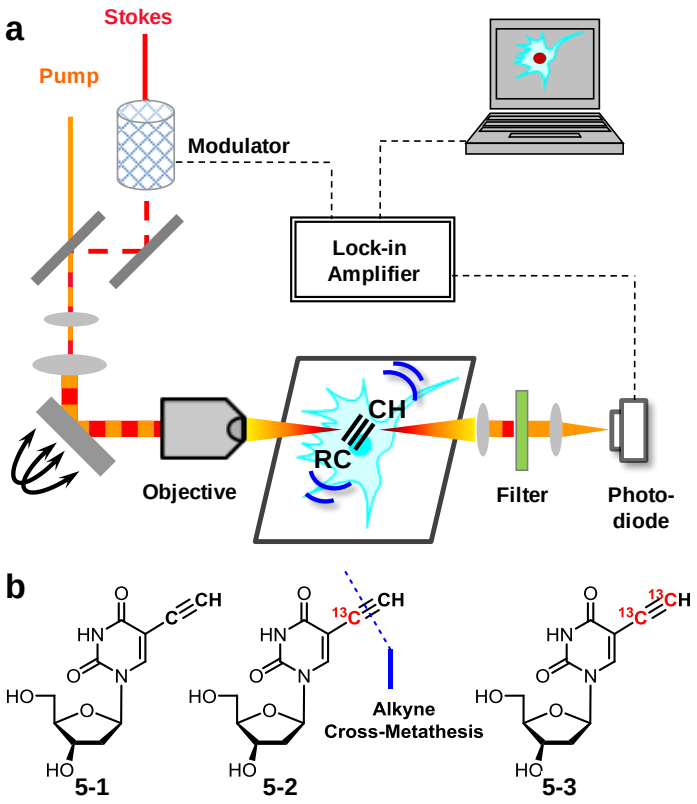



**Figure 5-1. Isotopically edited alkyne vibrational tags for chemical imaging by stimulated Raman scattering (SRS) microscopy**. (a) Setup of SRS microscope for alkyne vibrational imaging. When the energy difference between the pump and the Stokes photons matches with the alkyne vibration mode, their joint action will greatly accelerate the vibrational excitation of alkyne bonds. As a result of energy exchange between the input photons with the alkynes, the output pump and Stokes beams will experience intensity loss and intensity gain, respectively. Such intensity changes measured by SRS microscope generate concentration-dependent alkyne distributions in three-dimension (3D). (b) Structures of unlabeled, mono and bis $^{13}$C-labeled 5-ethynyl-2′-deoxyuridine (EdU). Mono-$^{13}$C-labeled EdU, **5-2**, is retrosynthetically disconnected using alkyne cross-metathesis chemistry.

## 5.3 Results

### 5.3.1 Theoretical consideration of Raman spectra

We first seek for some theoretical insights by using a simple classical mechanics model. The wavenumber of an alkyne stretching oscillation, $\bar{\nu}$, is related to the force

constant of the triple bond (k) as well as the mass of the two carbons ($m_1$ and $m_2$) using Hooke's law:

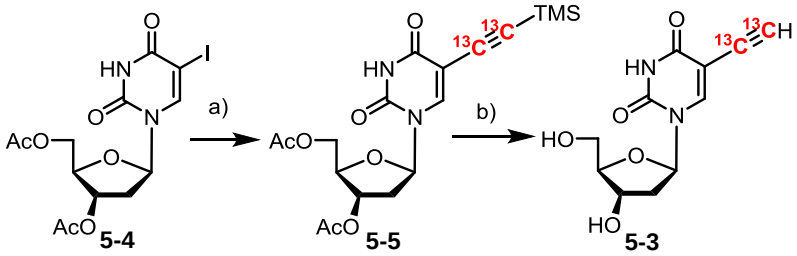
$$\bar{\nu} = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu}} \text{ , where } \mu = \frac{m_1 m_2}{m_1 + m_2} \qquad (5\text{-}1)$$

For a typical alkyne ($\bar{\nu}$ = 2125 $cm^{-1}$), if one of the alkyne carbons is substituted by $^{13}C$, the wavenumber of the stretching vibration is calculated to be 2084 $cm^{-1}$, assuming the change in the bond strength is negligible. Likewise, bis-$^{13}C$ labeled alkyne has a predicted wavenumber of 2042 $cm^{-1}$. Therefore, the expected spectral shifts of mono- and bis-isotopically labeled alkyne are 41 and 83 $cm^{-1}$, respectively.

In order to obtain a more accurate prediction, we further calculated the frequencies for the triple-bond stretching using density functional theory (DFT) at the B3LYP/6-31G* level of theory with the Scaled Quantum Mechanical force field method[24]. We use 5-ethynyl-2'-deoxyuridine (EdU, **5-1**) and its isotopologues, EdU-$^{13}C$ (**5-2**) and EdU-$^{13}C_2$ (**5-3**), as model compounds (Figure 5-1b). EdU is a thymidine analogue that incorporates into newly-synthesized DNA and is typically detected by fluorescent labeling via click chemistry[25]. We use EdU as a model alkynyl vibrational probe as it is the first reported alkyne-tagged molecule imaged with Raman microscopy[10]. The DFT-calculated vibrational frequencies for isotopically edited EdUs are 2126 $cm^{-1}$ (for **5-1**), 2076 $cm^{-1}$ (for **5-2**), and 2051 $cm^{-1}$ (for **5-3**). Given that the Raman peaks due to alkyne stretching are intrinsically sharp (typical FWHM = 14 $cm^{-1}$)[12], the $^{13}C$ isotope editing strategy should afford three spectroscopically resolvable alkyne vibrational tags for live cell Raman imaging and analysis.

### 5.3.2 Synthesis of isotopically edited EdUs

Encouraged by the theoretical predictions above, we developed chemical synthesis of the isotopically edited EdUs. Despite the commercial availability of EdU-$^{12}C_2$ (**5-1**), its isotopologues, EdU-$^{13}C$ (**5-2**) and EdU-$^{13}C_2$ (**5-3**), need to be chemically synthesized for characterization. EdU is prepared by alkynylation of 5-iododeoxyuridine by a Sonagashira coupling[26]. We used an analogous method in which acetylated 5-iododeoxyuridine (**5-4**) was subjected to Sonagashira coupling with trimethylsilylacetylene-$^{13}C_2$ to yield intermediate **5-5** (Scheme 5-1). Global deprotection with $K_2CO_3$ affords EdU-$^{13}C_2$ (**5-3**). However, the mono-$^{13}C$-labeled EdU (**5-2**) is a synthetic challenge due to the difficulty in creating the $^{13}C_1$ alkynyl building block.



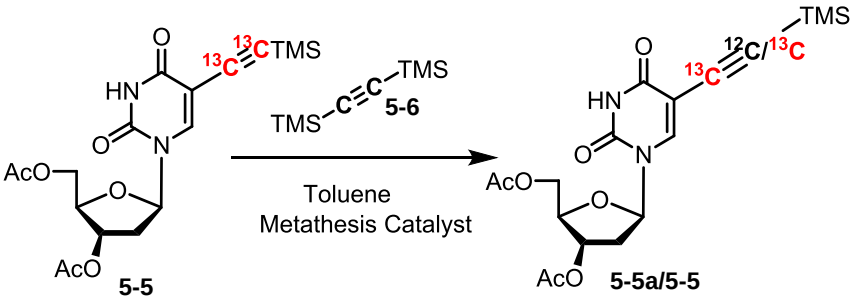

**Scheme 5-1. Synthesis of 5-3 by Sonagashira coupling.** Reagents and conditions: a) $Pd(OAc)_2$ (10% mol), $PPh_3$ (20% mol), CuI (10% mol), $Et_3N$ (3.0 eq.) and TMS$^{13}C\equiv^{13}CH$ (1.5 eq.), DMF, RT, 15 h, 72%; b) $K_2CO_3$ (5.0 eq.), MeOH/$H_2O$, RT, o/n, 75%.

We reasoned that an alkyne cross-metathesis disconnection would provide direct access to the mono-$^{13}C$-labeled alkyne from EdU-$^{13}C_2$ (**5-3**). Alkyne metathesis[27,28] is emerging as a viable tool for synthesis of complex molecules such as natural products and polymers. These advances are enabled by the development of new generations of catalysts that have higher catalytic activity and broader substrate compatibility[23]. Nevertheless, alkyne

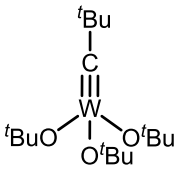
metathesis, especially cross-metathesis, had not been demonstrated on chemical biology reporters with a myriad of functional groups. Since metathesis of terminal alkynes remains a challenge[29], we turned to the TMS protected **5-5** as the substrate for metathesis[30]. It should be noted that TMS-protected alkynes exhibit reduced reactivity compared to typical internal alkynes[31].

We initially planned to metathesize **5-5** with a large excess (100 eq.) of bis(trimethylsilyl)acetylene, **5-6**, as the isotopic editing reagent to achieve a theoretical 99.5% conversion to mono-$^{13}$C-**5-5** at equilibrium. We first evaluated the commercially available Schrock catalyst[32](**5-7**) for alkyne metathesis (Table 5-1). Reacting **5-5** with 5 eq. **5-6** and 100 eq. **5-7** gave no conversion at room temperature (entry 1). When heated to 80 ℃, a side product was obtained,(entry 2, see supplementary information for detail), indicating that the glycosylamine was not compatible with the tungsten catalyst[33]. Therefore, we switched to a newly engineered podand-supported molybdenum catalyst (**5-8**) with high catalytic activity and stability developed by Zhang and Jyothish[34]. Substrate **5-5** is still inert to metathesis with **5-6** in the presence of catalyst **5-8** at room temperature (entry 3), but at 80 ℃ the reaction yields an inseparable 2.2 : 1 mixture of metathesis product and unreacted starting material, along with uncharacterized by-products (entry 4). Reducing the catalyst loading and the equivalence of **5-6** fails to yield **5-5** (entry 5), indicating that a high concentration of alkyne is required to drive the reaction of this relatively sterically encumbered substrate.

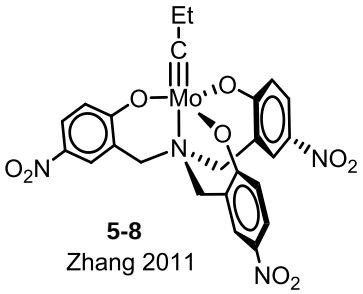
**Table 5-1. Investigation of conditions for the alkyne cross-metathesis**.




Catalysts:



**5-7**
Schrock 1981

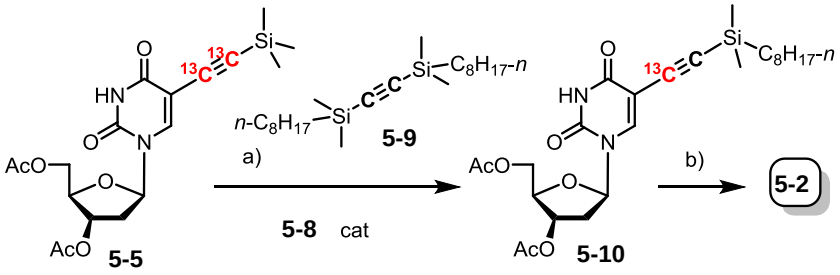


**5-8**
Zhang 2011

| Entry | Catalyst | **5-6** loading | Temp. | Time | Yield | **5-5a:5-5** |
|---|---|---|---|---|---|---|
| 1 | **5-7** , 5 eq. | 100 eq. | RT | 12h | N/R | |
| 2 | **5-7** , 5 eq. | 100 eq. | 80°C | 12h | Side Product | |
| 3 | **5-8** , 5 eq. | 100 eq. | RT | 12h | N/R | |
| 4 | **5-8** , 5 eq. | 100 eq. | 80°C | 12h | 20% | 2.2 : 1 |
| 5 | **5-8** , 0.5 eq. | 10 eq. | 80°C | 12h | N/R | |

Based on these results, we decided not to search for another protocol to push the conversion of the bulky substrate to >95%. Instead, we developed an alternative strategy to separate the mono-$^{13}$C product from unreacted starting material (Scheme 5-2). We reason that a different protecting group from TMS, once introduced to the product, would render the product isolable by chromatography. We chose octyldimethylsilyl as it is more hydrophobic than the TMS group but not too bulky to inhibit the reaction. We prepared bis(octyldimethylsilyl)acetylene (**5-9**) and metathesized **5-5** with **5-9** in the presence of **5-8**. To our delight, the desired product, compound **5-10,** was isolated in pure form in a moderate

yield (33% based on recovered starting materials). Deprotection with $K_2CO_3$ and TBAF afforded the final product, EdU-$^{13}$C (**5-2**). The isotopic features of **5-2** and **5-3** are characterized by mass spectrometry, $^1$HNMR and by the coupling between $^{13}$C and $^1$H (See supplementary information section for detail).




**Scheme 5-2. Synthesis of 5-2 by alkyne cross-metathesis.** Reagents and conditions: a) **5-9** (100 eq.), **5-8** (5 eq.), $CCl_4$, 70 °C, 8 h, 27%, 33% B.R.S.M.; b) $K_2CO_3$ , TBAF, MeOH/$H_2O$, RT, 7 h, 50%.

### 5.3.3 Characterization of multi-color alkyne vibrational tags by Raman spectroscopy and SRS imaging

With compounds **5-1**, **5-2** and **5-3** in hand, we first measured their Raman spectroscopic properties as labeling reagents for *de novo* DNA synthesis in proliferating cells. HeLa cells were treated with 0.1 mM **5-1**, **5-2**, and **5-3**, respectively**,** for 15 hr, then fixed and analyzed using a Raman micro-spectrometer. The three isotopologues are all biochemically active, as alkyne stretching peaks are detectable in the cell nuclei of all three samples (Figure 5-2). Incorporated EdU (**5-1**) exhibits a Raman peak at 2125 cm$^{-1}$ as expected, while incorporated EdU-$^{13}$C (**5-2**) and EdU-$^{13}C_2$ (**5-3**) have displayed markedly shifted peaks at 2077 and 2048 cm$^{-1}$, respectively. Therefore, the central wavenumbers of the measured new peaks are in good agreement with the DFT predictions (2076 cm$^{-1}$ and 2051 cm$^{-1}$). More importantly, thanks to the intrinsic sharpness (typical FWHM = 14 cm$^{-1}$)

of alkyne peaks, the spectral shifts of EdU-$^{13}$C (**5-2**) and EdU-$^{13}C_2$ (**5-3**) are just large enough so that all three peaks are completely resolved from each other.

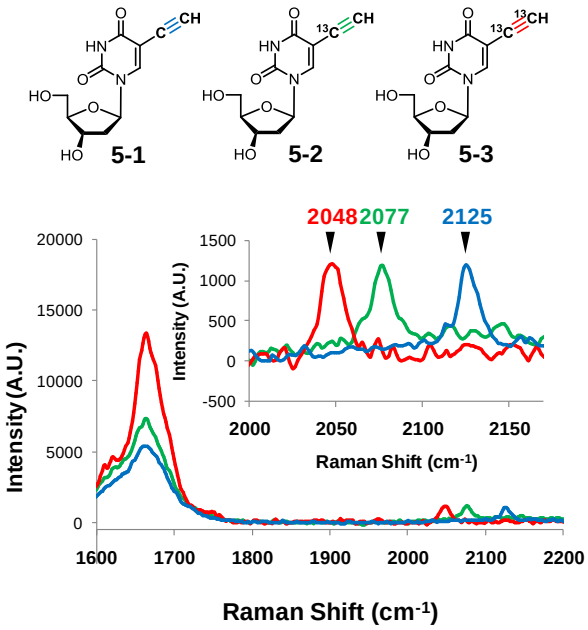



**Figure 5-2. Raman spectra of HeLa cells incubated with three isotopically edited EdUs.** Spectra are acquired from nucleus region of fixed cells after incubated with either **5-1**, **5-2** or **5-3**. Amide bond stretchings at 1655 cm$^{-1}$ are shown as reference. The spectra are normalized according to the alkyne peak. Inset: enlarged Raman spectra from 2000 cm$^{-1}$ to 2170 cm$^{-1}$.

We next evaluated compound **5-1**, **5-2** and **5-3** in live mammalian cells under SRS microscopy (Figure 5-3). In our SRS microscope setup, the pulse widths of both pump and Stokes lasers are set to 6 ps, corresponding to an excitation profile of 6 cm$^{-1}$. This excitation profile is slightly narrower than 14 cm$^{-1}$ (FWHM) of alkyne peaks, rendering both an efficient and a selective SRS excitation to the specifically labeled alkynes[12]. To test the orthogonality of **5-1**, **5-2** and **5-3**, live HeLa cells treated with each probe were imaged under all five channels: 1655 cm$^{-1}$ (amide channel, total protein), 2000 cm$^{-1}$ (off-resonance channel), 2048 cm$^{-1}$, 2077 cm$^{-1}$ and 2125 cm$^{-1}$ (on-resonance channels with the three alkyne

peaks, respectively). Images are acquired sequentially under the same laser powers. As shown in Figure 5-3, for EdU (**5-1**), a prominent signal is observed at the 2125 $cm^{-1}$ on-resonance channel, depicting the newly synthesized DNA inside nucleus. In contrast, the 2000 $cm^{-1}$, 2048 $cm^{-1}$ and 2077 $cm^{-1}$ channels detect only weak and dispersive background which may be attributed to optical cross-phase modulations[35]. Similarly, when using EdU-$^{13}C$ (**5-2**) or EdU-$^{13}C_2$ (**5-3**) as the probe, DNA synthesis signals are only detected at the corresponding on-resonance channel of 2077 $cm^{-1}$ or 2048 $cm^{-1}$ (Figure 5-3). These SRS images, along with the spectroscopic data, unambiguously prove isotopic editing as a viable strategy of spectral shifting for multi-channel Raman study of alkyne-tagged molecules.

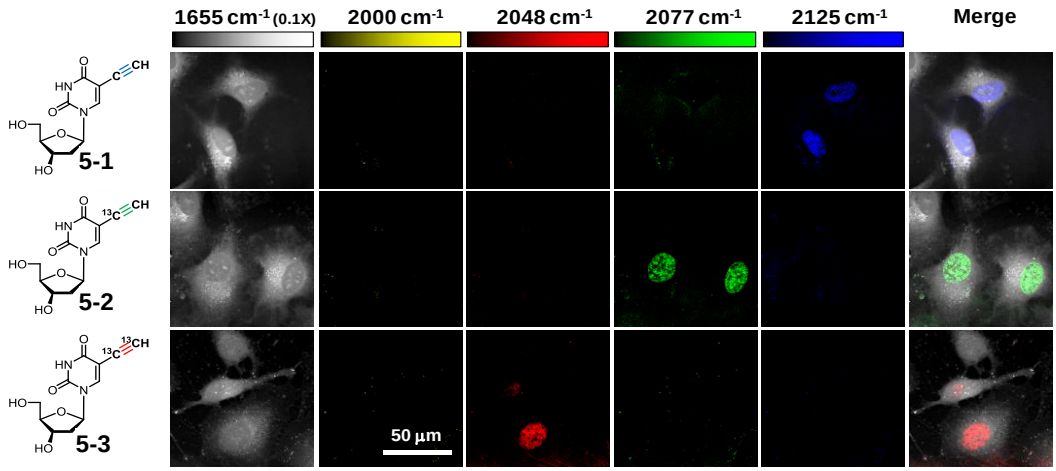



**Figure 5-3. Live cell SRS imaging of DNA synthesis in HeLa cells incubated with isotopically-edited EdUs.** For each sample incubated with either **5-1** or **5-2** or **5-3**, images are acquired in 5 different Raman channels: 1655 $cm^{-1}$ (amide bond), 2000 $cm^{-1}$ (off-resonant), 2048 $cm^{-1}$ (on-resonant with **5-3**), 2077 $cm^{-1}$ (on-resonant with **5-2**) and 2125 $cm^{-1}$ (on-resonant with **5-1**) in sequential mode. Images are acquired in 512 X 512 pixels with a pixel dwell time of 40 µs.

### 5.3.4 Three-color SRS imaging of isotopic alkyne vibrational tags

Currently, there exists a variety of alkyne derivatized metabolic labeling reagents for studying DNA, RNA, proteins, lipids and other bio-molecules[36]. These probes were developed originally for the subsequent Cu-catalyzed click reactions[37] but have proven to be also suitable for direct Raman imaging[12]. We choose three well-documented alkynyl reporters for our demonstration: ethylnyluridine (EU, **5-11**) as a small-molecule precursor for RNA synthesis[38], 17-octadecynoic acid (17-ODYA, **5-12**) as a lipid precursor as well as a protein post-translational modification reagent[39] and EdU (**5-1**) for DNA synthesis. As shown in Figure 4a, the Raman spectrum of EU (**5-11**, aqueous solution), EdU (**5-1**, aqueous solution) and 17-ODYA (**5-12**, neat) exhibit peaks at 2126, 2123 and 2120 $cm^{-1}$, respectively. The near complete spectral overlap among the three Raman peaks is thus prohibitive for further attempts of multi-color imaging.

Enlightened by the isotopic editing concept, we used the mono-$^{13}C$ labeled EdU-$^{13}C$ (**5-2**), a bis-$^{13}C$ labeled EU-$^{13}C_2$ (**5-13**) and 17-ODYA (**5-12**) to address this issue. We first verified the Raman spectra of the isotopically edited probes as shown in Figure 5-4b. In aqueous solutions, EU-$^{13}C_2$ (**5-13**), has a shifted peak at 2053 $cm^{-1}$, which is well-resolved from the 2077 $cm^{-1}$ peak of EdU-$^{13}C$ (**5-2**, aqueous solution) and the 2120 $cm^{-1}$ peak of 17-ODYA (**5-12**). Therefore, we moved on to test the three-color live cell SRS imaging of RNA, DNA and lipid metabolism using EU-$^{13}C_2$ (**5-13**), EdU-$^{13}C$ (**5-2**) and 17-ODYA (**5-12**).

HeLa cells were treated with **5-13**, **5-2**, and **5-12** simultaneously before being imaged by SRS microscopy. Two living cells are captured in Figure 5-4c. In the 2053 $cm^{-1}$ channel, the signal for total RNA is observed mainly inside the nucleus of both cells, with intense signal at nucleoli where ribosomal RNA assembly takes place. In the 2077 $cm^{-1}$ channel, the

EdU-$^{13}$C signal shows a nuclear distribution of newly synthesized DNA in one cell but not the other (indicating different cell cycle status). In the 2125 $cm^{-1}$ channel, the signal is derived from 17-ODYA (**5-12**) incorporation to lipids that exhibit cytosolic localization. These observations are in accordance with the known cell biology that DNA replication happens only when cells are progressing through the S phase while fatty acid take-up and RNA synthesis processes are less dependent on cell cycles. Little background is observed in the 2000 $cm^{-1}$ off-resonance channel. Amide channel at 1655 $cm^{-1}$, interpreted as total protein signal, is shown as a reference. In the final merged image of Figure 5-4 c, three alkynyl molecules could be differentiated unambiguously, allowing for multiplex studies of co-localization and interactions.

Overall, by introducing isotopically edited alkyne tags we have successfully rendered different alkynyl metabolic probes three distinctive Raman "colors". This approach could be readily applied to studying other combinations of small-molecule reporters. It is noteworthy that, the live-cell three-color Raman imaging of isotopic alkyne-tagged metabolic reporters demonstrated here would be difficult to achieve otherwise, for example, by using fluorescent reporters due to the limited choices of bioorthogonal chemistry.

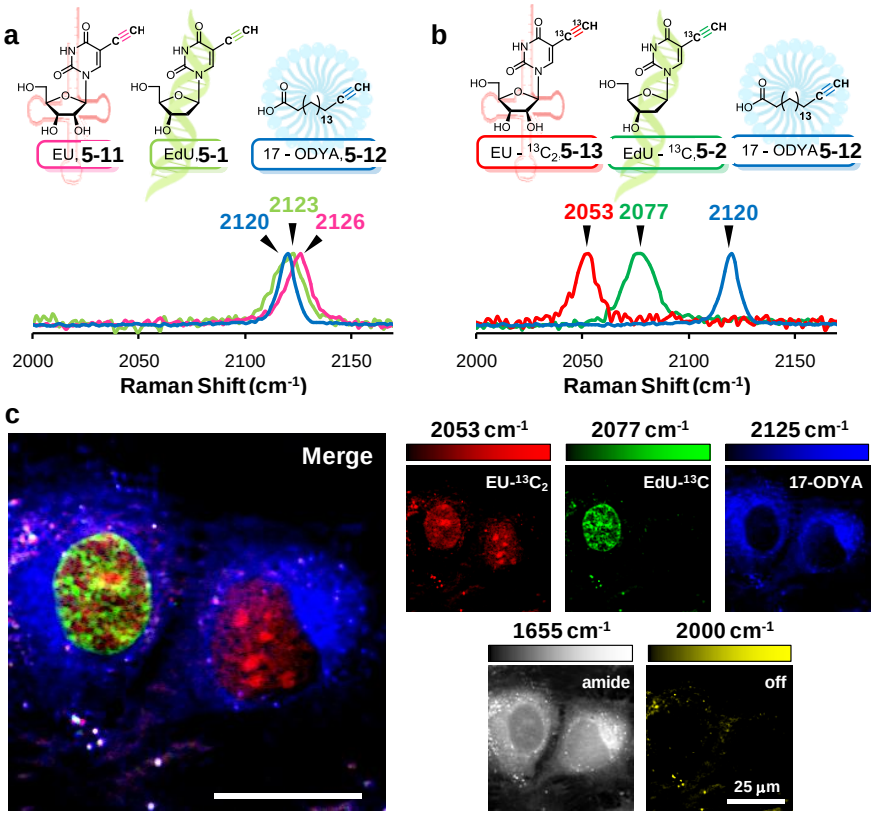



**Figure 5-4. Three-color chemical imaging using isotopically edited alkyne tags.** (a) Structures and normalized Raman spectra of RNA probe EU (**5-11**), DNA probe EdU (**5-1**), fatty acid probe 17-octadecynoic acid (**5-12**). (b) Structures and normalized Raman spectra of isotopically edited EU-$^{13}C_2$ (**5-13**), EdU-$^{13}C$ (**5-2**) and 17-ODYA (**5-12**). (c) Three-color SRS imaging of nascent RNA, DNA and fatty acyl derivatives in live HeLa cells by spectral targeting of different isotopically edited alkyne tags. Images are acquired in 5 different Raman channels: 1655 cm$^{-1}$ (amide bond), 2000 cm$^{-1}$ (off-resonant), 2053 cm$^{-1}$ (on-resonant with **5-13**), 2077 cm$^{-1}$ (on-resonant with **5-2**) and 2125 cm$^{-1}$ (on-resonant with **5-12**) in sequential mode. Images are acquired in 341 X 341 pixels with a pixel dwell time of 40 μs.

## 5.4 Discussion

Our spectroscopic and imaging studies demonstrate that the mono and bis- $^{13}C$ isotopically barcoded alkyne tags can shift Raman peaks to be well resolved from their unmodified counterpart. This isotopic editing strategy has enabled us to perform live cell

three-color chemical imaging of DNA, RNA and fatty-acid metabolism simultaneously, by three ethynyl-derivatized reporters: EU-$^{13}C_2$, EdU-$^{13}C$ and 17-ODYA. Impressively, EdU-$^{13}C$ was prepared via an alkyne cross-metathesis reaction using a podand-supported molybdenum(VI) catalyst[34]. Alkyne vibrational tags can now be employed in three mutually orthogonal versions as reporters for high-resolution, multi-color chemical imaging and subsequent studies on spatial co-localization and functional interactions. This work thus represents important progress toward chemical imaging of complex biological processes in live cells.

From a synthetic chemistry point of view, this work is not only a rare demonstration of preparing a bio-probe using alkyne metathesis, but also one of the few examples in which relatively bulky (trialkylsilyl)alkynes are subjected to cross-metathesis. Notably, the molybdenum catalyst used here is mild enough for the sensitive hemiaminalether structure of EdU. Considering that the metathesis of TMS-protected alkyne was proposed[30] in 1983 and first demonstrated[31] in 2001, it is clear that the evolution of catalysts has been constantly broadening the scope of alkyne metathesis. We expect that, with the development of milder and more active catalysts[40,41], alkyne metathesis would be carried out under mild conditions with higher yields. As progress has been made on metathesizing unprotected terminal alkynes[29] and on alkyne metathesis in protic media[42], the work reported here will be regarded not only as an early effort of using alkyne metathesis in the preparation of bioimaging reagent, but also as a prelude of alkyne metathesis as a general tool in bio-conjugation and bio-orthogonal chemistry[43].

From an isotopic chemistry point of view, this work reaffirms the unique role of isotopes in vibrational spectroscopy. While fluorescent dyes and proteins change their color

by modulating the delocalization of conjugated π electrons and quantum dots by quantum confinement effects, alkyne vibrational palette harness the fundamental mass-energy relation to tune the vibrational frequency. Indeed, isotopic editing has been broadly applied in chemistry studies including multi-color Raman imaging of $^{13}$C-doped carbon nanotubes[17], studying β-sheet structure with $^{13}$C labeled carbonyls as infrared (IR) probes[18], characterizing single-molecule surface-enhanced Raman spectroscopy[19,20]and tuning spectroscopic profiles of environmentally sensitive IR probes[21,22]. By overcoming the synthetic hurdle via alkyne metathesis, here we have achieved the isotopic editing of alkynes. It is also noteworthy that two-color alkyne imaging could be occasionally achieved in rare cases when molecules have resolvable Raman shifts (e.g., propargylcholine have a shifted Raman peak presumably due to a nearby positive charge[12] or by introducing other bulkier groups such as diynes[11]). Isotopic editing, however, offers a general and modular way to change Raman peaks of alkynyl molecules. On a broader perspective, the isotopically edited alkynyl molecules could be further used in modalities beyond vibrational imaging, like MRI[44] and mass-spec based imaging/analysis[45].

From a Raman imaging instrumentation point of view, the SRS images in this work are recorded at a pixel dwell time of 40 μs, ~4500 times faster than in a spontaneous Raman imaging setup[10]. Nevertheless, isotopically edited alkynes should also be applicable to more accessible spontaneous Raman microscopes, which requires longer time to reconstruct an image but offers a full spectrum at each pixel[46,47]. Moreover, the SRS microscope setup in this work could be further upgraded towards advanced multi-color imaging applications with alkyne vibrational tags. First, higher stimulated Raman loss signal could be achieved by equipping stronger lasers, i.e., fiber lasers, rendering more sensitive detections of alkynyl

molecules[48]. Second, the background signal derived from cross-phase modulation could be reduced by a recently reported spectral modulation technique[49]. Lastly, the newly developed rapidly tunable optical parametric oscillator would enable line-by-line acquisition of multi-color SRS images, which would reduce the potential artifacts caused by sample motions in current frame-by-frame imaging acquisition mode[50].

From an imaging probe point of view, the alkyne vibrational palette described here adds to a new dimension of imaging reagents for studying metabolically labeled biomolecules as well as non-immobilized cellular small molecules. In comparison, fluorescence-based methods of imaging metabolically labeled biomolecules largely rely on Cu-catalyzed click-chemistry on fixed samples[36]. The development of Cu-free click-chemistry is, while fast[51–53], still limited to few functional groups such as cyclooctyne[54] and tetrazine[55]. Therefore, the alkyne vibrational tags, now available in three colors, represent a major advantage. Finally, the isotopic alkyne-tagged molecules would retain their click-chemistry reactivity with subsequent biochemical pull-down reagents, rendering a combination of complementary methods on different levels.

## 5.5 Conclusion

We present the isotopic editing of alkyne vibrational tags: the theory-guided design, chemical synthesis via alkyne metathesis, spectroscopic characterization and their application in three-color live-cell chemical imaging using SRS microscopy. The alkyne vibrational palette offers a new dimension to multi-color chemical imaging, complementing fluorescence microscopy in multiplex studies for visualization, localization and interaction of a variety of biochemical species in live cells. Moreover, the strategic application of alkyne

cross-metathesis in the chemical synthesis of isotopic barcodes echoes the impact of emerging synthetic methodology to the development of chemical reporters for burgeoning bio-imaging applications.

## 5.6 Methods

### 5.6.1 Stimulated Raman Scattering microscopy

The microscopy set up was previously described[12]. Briefly, an integrated light source (picoEMERALD with custom modification, Applied Physics & Electronics, Inc), consisting of a Stokes beam (1064 nm, 6 ps, 80 MHz repetition rate) modulated at 8 MHz and a spatially- temporally overlapping pump beam (tunable, 720 to 990 nm, 5-6 ps, 80 MHz repetition rate), is coupled into an inverted multi-photon laser-scanning microscope (FV1200MPE, Olympus). Lasers are delivered to the cell samples through a 60X water objective (UPlanAPO/IR, 1.2 N.A., Olympus) and then collected with a condenser lens (*oil immersion, 1.4* N.A., Olympus). The Stokes beam is then blocked with a high O.D. bandpass filter (890/220 CARS, Chroma Technology) while the pump beam is imaged onto a Si photodiode (FDS1010, Thorlabs). To detect the stimulated Raman loss, output current from the Si photodiode is terminated, filtered, and demodulated using a lock-in amplifier (SR844; Stanford Research Systems). The output of the lock-in amplifier is sent to the microscope through an analog interface box (FV10-ANALOG, Olympus) and images are reconstructed using Fluoview software (Olympus). The imaging experiments in this study are performed with 168 mW pump beam and 134 mW Stokes beam (power measured after objective). Images are acquired with a 40 μs pixel dwell time and a 10 μs time constant from the lock-

in amplifier. For multi-channel SRS experiments, images are acquired in sequential mode with a laser tuning duration of 40 s – 80 s between channels. The total acquisition time of a 5-channel SRS image is 5 min.

### 5.6.2 Raman Spectroscopy

Raman spectra were acquired at room temperature with a Raman spectrometer (inVia Raman microscope; Renishaw) equipped with a 532 nm diode laser through a 50×, 0.75 N.A. objective (NPLAN EPI; Leica). Spectra were acquired in 100 s and processed using WiRE software.

### 5.6.3 Live cell imaging

HeLa cells were seeded on glass coverslips and cultured in DMEM with 10% (v/v) FBS and 1% penicillin/streptomycin (Invitrogen). For EdU imaging experiments, media were changed to FBS-free DMEM for 24 h to synchronize cell cycles. The media were then changed back to DMEM with 10% FBS at the time isotopically edited EdUs were added to media to a final concentration of 100 μM. 15 h later, the coverslips were washed with PBS and assembled into imaging chambers using imaging spacers (GBL 654008, Sigma) filled with PBS. For three-color imaging with EdU-$^{13}$C, EU-$^{13}C_2$ and 17-ODYA, HeLa cells were synchronized by changing to FBS-free DMEM for 24 h, followed by incubating with DMEM (10% FBS) in the presence of 100 μM EdU-$^{13}$C and 50 μM 17-ODYA. 9 h later, EU-$^{13}C_2$ was also added to the media (final concentration 2 mM). The cells were incubated for another 6 h before washed with PBS and assembled into imaging chambers for SRS imaging.

## 5.7 Supporting Information

### 5.7.1 Density Functional Theory Calculation Methods

DFT calculations were done using version 8.3 of Jaguar, Schrödinger, Inc.'s electronic structure software. All calculations (geometry optimizations and frequency calculations) were done in water (isolv=2) using the B3LYP functional and the 6-31G* basis set. Final geometries of structures **5-1-3** follow. For the frequency calculations (ifreq=1), ultrafine grids were used (gdftder2=gdftcphf=-13), and Peter Pulay et. al.'s Scaled Quantum Mechanical force field method was applied (isqm=1), which scales individual elements of the Hessian according to the type of vibration. This corrects for systematic deficiencies in the wavenumbers calculated at the B3LYP/6-31G* level of theory (Baker, J. et. al., *J. Phys. Chem. A*, **1998**, *102*, 1412–1424).

Structure **5-1** final geometry:

| | | | |
|---|---|---|---|
| O1 | -2.8843370000 | 0.8575900000 | -3.4793340000 |
| C2 | -1.8003270000 | 1.4263520000 | -2.7421970000 |
| C3 | -1.2288950000 | 0.3621640000 | -1.7946190000 |
| O4 | -0.0249810000 | 0.9317150000 | -1.2096030000 |
| C5 | -2.1424180000 | -0.0156100000 | -0.6094230000 |
| C6 | -0.1841170000 | 1.2291760000 | 0.1581130000 |
| O7 | -1.9606130000 | -1.3759050000 | -0.1987870000 |
| C8 | -1.6522190000 | 0.9204780000 | 0.5098040000 |
| N9 | 0.8221990000 | 0.4623110000 | 0.9809170000 |
| C10 | 1.2068510000 | -0.8050310000 | 0.6606400000 |
| C11 | 1.3520580000 | 1.1289880000 | 2.0892210000 |
| C12 | 2.1420640000 | -1.5072390000 | 1.3753880000 |
| O13 | 1.0011010000 | 2.2607050000 | 2.4165530000 |
| N14 | 2.3096070000 | 0.4209620000 | 2.7890800000 |
| C15 | 2.5156730000 | -2.8336060000 | 1.0119400000 |
| C16 | 2.7758650000 | -0.8723740000 | 2.5357420000 |
| C17 | 2.8218410000 | -3.9574600000 | 0.6763290000 |
| O18 | 3.6365040000 | -1.3807960000 | 3.2516210000 |
| H19 | -3.1726400000 | 1.5236800000 | -4.1363840000 |
| H20 | -0.9932000000 | 1.7595850000 | -3.4102870000 |
| H21 | -2.1283920000 | 2.2995720000 | -2.1546690000 |
| H22 | -0.9458380000 | -0.5338460000 | -2.3616490000 |
| H23 | -3.1945960000 | 0.1737000000 | -0.8542380000 |
| H24 | 0.0846290000 | 2.2710990000 | 0.3316300000 |
| H25 | -2.2817100000 | -1.9495040000 | -0.9247400000 |
| H26 | -1.7702930000 | 0.4796560000 | 1.5036960000 |
| H27 | -2.2191720000 | 1.8568770000 | 0.4841150000 |
| H28 | 0.7241700000 | -1.2327500000 | -0.2104570000 |
| H29 | 2.7154710000 | 0.9086180000 | 3.5924050000 |
| H30 | 3.0950010000 | -4.9516330000 | 0.3922090000 |

Structure **5-2** final geometry (atom C15 was set to the isotope C-13):

| | | | |
|---|---|---|---|
| O1 | -2.8841450000 | 0.8579240000 | -3.4802570000 |
| C2 | -1.8001350000 | 1.4266860000 | -2.7431200000 |
| C3 | -1.2287030000 | 0.3624980000 | -1.7955420000 |
| O4 | -0.0247890000 | 0.9320490000 | -1.2105260000 |
| C5 | -2.1422260000 | -0.0152760000 | -0.6103460000 |
| C6 | -0.1839250000 | 1.2295100000 | 0.1571900000 |
| O7 | -1.9604210000 | -1.3755710000 | -0.1997100000 |
| C8 | -1.6520270000 | 0.9208120000 | 0.5088810000 |
| N9 | 0.8223910000 | 0.4626450000 | 0.9799940000 |
| C10 | 1.2070430000 | -0.8046970000 | 0.6597180000 |
| C11 | 1.3522500000 | 1.1293220000 | 2.0882980000 |
| C12 | 2.1422560000 | -1.5069050000 | 1.3744650000 |
| O13 | 1.0012930000 | 2.2610390000 | 2.4156300000 |
| N14 | 2.3097990000 | 0.4212960000 | 2.7881570000 |
| C15 | 2.5158650000 | -2.8332720000 | 1.0110170000 |
| C16 | 2.7760570000 | -0.8720400000 | 2.5348190000 |
| C17 | 2.8220330000 | -3.9571260000 | 0.6754060000 |
| O18 | 3.6366950000 | -1.3804620000 | 3.2506990000 |
| H19 | -3.1724480000 | 1.5240140000 | -4.1373060000 |
| H20 | -0.9930080000 | 1.7599190000 | -3.4112090000 |
| H21 | -2.1282000000 | 2.2999060000 | -2.1555910000 |
| H22 | -0.9456460000 | -0.5335130000 | -2.3625720000 |
| H23 | -3.1944040000 | 0.1740340000 | -0.8551610000 |
| H24 | 0.0848210000 | 2.2714330000 | 0.3307080000 |
| H25 | -2.2815180000 | -1.9491700000 | -0.9256620000 |
| H26 | -1.7701010000 | 0.4799900000 | 1.5027730000 |
| H27 | -2.2189800000 | 1.8572110000 | 0.4831930000 |
| H28 | 0.7243620000 | -1.2324160000 | -0.2113800000 |
| H29 | 2.7156630000 | 0.9089520000 | 3.5914820000 |
| H30 | 3.0951930000 | -4.9512990000 | 0.3912860000 |

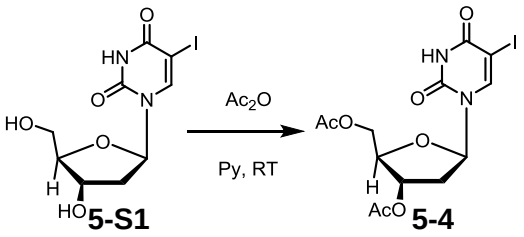
Structure **5-3** final geometry (atoms C15 and C17 were set to the isotope C-13):

| | | | |
|---|---|---|---|
| O1 | -2.8837710000 | 0.8584020000 | -3.4815580000 |
| C2 | -1.7997610000 | 1.4271640000 | -2.7444210000 |
| C3 | -1.2283290000 | 0.3629760000 | -1.7968430000 |
| O4 | -0.0244150000 | 0.9325270000 | -1.2118270000 |
| C5 | -2.1418520000 | -0.0147980000 | -0.6116470000 |
| C6 | -0.1835500000 | 1.2299880000 | 0.1558890000 |
| O7 | -1.9600470000 | -1.3750930000 | -0.2010110000 |
| C8 | -1.6516530000 | 0.9212900000 | 0.5075800000 |
| N9 | 0.8227650000 | 0.4631230000 | 0.9786930000 |
| C10 | 1.2074170000 | -0.8042190000 | 0.6584170000 |
| C11 | 1.3526240000 | 1.1298000000 | 2.0869970000 |
| C12 | 2.1426300000 | -1.5064270000 | 1.3731640000 |
| O13 | 1.0016670000 | 2.2615170000 | 2.4143290000 |
| N14 | 2.3101730000 | 0.4217740000 | 2.7868560000 |
| C15 | 2.5162390000 | -2.8327940000 | 1.0097160000 |
| C16 | 2.7764310000 | -0.8715620000 | 2.5335180000 |
| C17 | 2.8224070000 | -3.9566480000 | 0.6741050000 |
| O18 | 3.6370700000 | -1.3799840000 | 3.2493980000 |
| H19 | -3.1720740000 | 1.5244920000 | -4.1386080000 |
| H20 | -0.9926340000 | 1.7603970000 | -3.4125100000 |
| H21 | -2.1278250000 | 2.3003840000 | -2.1568920000 |
| H22 | -0.9452720000 | -0.5330340000 | -2.3638730000 |
| H23 | -3.1940300000 | 0.1745120000 | -0.8564620000 |
| H24 | 0.0851950000 | 2.2719110000 | 0.3294070000 |
| H25 | -2.2811440000 | -1.9486920000 | -0.9269630000 |
| H26 | -1.7697270000 | 0.4804680000 | 1.5014720000 |
| H27 | -2.2186060000 | 1.8576890000 | 0.4818920000 |
| H28 | 0.7247360000 | -1.2319380000 | -0.2126810000 |
| H29 | 2.7160370000 | 0.9094300000 | 3.5901810000 |
| H30 | 3.0955670000 | -4.9508210000 | 0.3899850000 |

### 5.7.2 Synthetic Chemistry

**General Methods.** Unless otherwise noted, reagents and solvents were obtained from Sigma-Aldrich, Alfa-Aesar and Cambridge Isotopes and were used without further purification. Flash chromatography was performed on a Teledyne ISCO CombiFlashRf using RediSepRf silica gel columns. Nuclear magnetic resonance (NMR) spectra were recorded on a Bruker 500 (500MHz) Bruker 400 (400 MHz) or Bruker 300 (300 MHz) Fourier Transform (FT) NMR spectrometers at Columbia University, Chemistry Department. $^1$H NMR spectra are tabulated in the following order: multiplicity (s, singlet; d, doublet; t, triplet; m, multiplet; br, broad), number of protons. Chemical shifts are referenced to the solvent residual peak. Fast Atom Bombardment (FAB) high resolution mass spectra (HRMS) were recorded on a JMS-HX110A mass spectrometer, and low resolution electron spray ionization (ESI) MS were recorded on a JMS-LC mate mass spectrometer.

Synthesis of **5-4**:



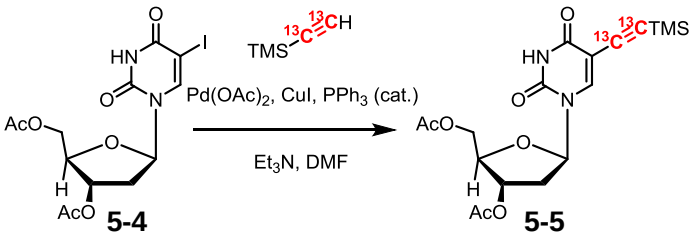

To a solution of 5-iodo-2'-deoxyuridine (**5-S1**, 150 mg, 0.42 mmol) in 1.5 ml of pyridine was added 0.4 ml (0.42 mmol) acetic anhydride at 0 ℃. The resulting mixture was warmed up to room temperature and stirred for 4 h, then poured into 5 ml of cold 1 N $NaHSO_4$ and extracted with ethyl acetate three times. The organic layer was washed with saturated $NaHCO_3$ and brine, dried over anhydrous $Na_2SO_4$ and concentrated. The crude

product was purified by column chromatography on silica gel (0-70% Ethyl acetate in Hexanes) to give **5-4** (157.3 mg, 0.36 mmol, 85%) as a white solid.

$^1$H NMR (400 MHz, $CDCl_3$) δ ppm: 8.46 (s, 1 H), 7.97 (s, 1 H), 6.28 (dd, *J* = 8.2, 5.7 Hz, 1 H), 5.27 – 5.19 (m, 1 H), 4.41 (dd, *J* = 12.3, 3.2 Hz, 1 H), 4.34 (dd, *J* = 12.3, 2.9 Hz, 1 H), 4.30 (q, *J* = 2.9 Hz, 1 H), 2.54 (ddd, *J* = 14.3, 5.7, 2.1 Hz, 1 H), 2.21 (s, 3 H), 2.20 – 2.13 (m, 1 H), 2.12 (s, 3 H).

MS (APCI+) m/z Calcd. for $C_{13}H_{16}IN_2O_7$ $[M+H]^+$: 439.0. Found: 438.8

Synthesis of **5-5**:



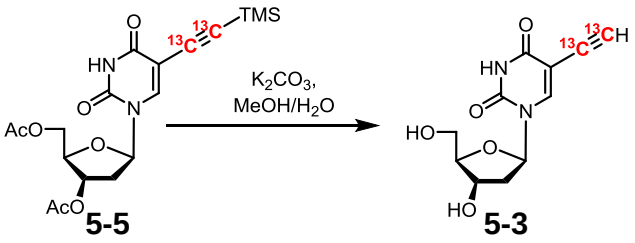

To an oven-dried vial was added **5-4** (72 mg, 164 μmol), $Pd(OAc)_2$ (3.6 mg, 16 μmol), $PPh_3$ (8.6 mg, 33 μmol), CuI (3.1 mg, 16 μmol), DMF (2 ml), $Et_3N$ (50 mg, 69 μl, 492 μmol) and TMS$^{13}$C≡$^{13}$CH (25 mg, 250 μmol) under Ar. The yellow mixture was stirred at RT for 15 h before concentrated *in vacuo*. The residue was purified by column chromatography on silica gel (0-70% Ethyl acetate in Hexanes) to give **5-5** (48.4 mg, 118 μmol, 72%) as a thin film.

$^1$H NMR (400 MHz, Methanol-$d_4$) δ 7.98 (d, *J* = 5.0 Hz, 1 H), 6.23 (dd, *J* = 7.8, 6.0 Hz, 1 H), 5.28 (dt, *J* = 6.7, 2.6 Hz, 1 H), 4.36 (t, *J* = 3.1 Hz, 2 H), 4.34 – 4.28 (m, 1 H), 2.50 (ddd, *J* = 14.5, 6.0, 2.5 Hz, 1 H), 2.39 (ddd, *J* = 14.5, 7.9, 6.6 Hz, 1 H), 2.16 (s, 3 H), 2.09

(s, 3 H), 0.20 (d, *J* = 2.5 Hz, 9 H). $^{13}$C NMR (101 MHz, $CD_3OD$) δ 99.54 (d, *J* = 140.5 Hz), 96.95 (d, *J* = 140.5 Hz).

MS (APCI+) m/z Calcd. for $C_{16}{}^{13}C_2H_{25}N_2O_7Si$ $[M+H]^+$: 411.2. Found: 411.0

Synthesis of **3**:




To a solution of **5-5** (3.5 mg, 8.5 μmol) in 0.9 ml MeOH and 0.1 ml $H_2O$ was added $K_2CO_3$ (6.0 mg, 43 μmol) at RT. The reaction was stirred overnight before concentrated *in vacuo*. The residue was purified by reverse phase HPLC to give compound **5-3** (1.6 mg, 6.4 μmol, 75%) as a thin film.

HPLC condition: 20 min gradient elution using $H_2O$:MeCN starting from 100:0 to 85:15. Retention time: 15.4 min

$^1$H NMR (400 MHz, $CD_3OD$) δ: 8.39 (d, *J* = 5.6 Hz, 1 H); 6.24 (t, *J* = 6.4 Hz, 1 H); 4.40 (m, 1 H); 3.94 (dd, *J* = 6.4, 3.2 Hz, 1 H); 3.82 (dd, *J* = 12, 3.2 Hz, 1 H); 3.73 (dd, *J* = 12, 3.6 Hz, 1 H); 3.53 (dd, *J* = 250.4, 54.8 Hz, 1 H); 2.32 (ddd, *J* = 13.6, 6, 3.6 Hz, 1 H); 2.23 (m, 1 H). $^{13}$C NMR (101 MHz, $CD_3OD$) δ 82.87 (d, *J* = 180.4 Hz), 75.85 (d, *J* = 180.3 Hz).

MS (FAB+) m/z Calcd. for $C_9{}^{13}C_2H_{13}N_2O_5$ $[M+H]^+$: 255.09. Found: 255.11

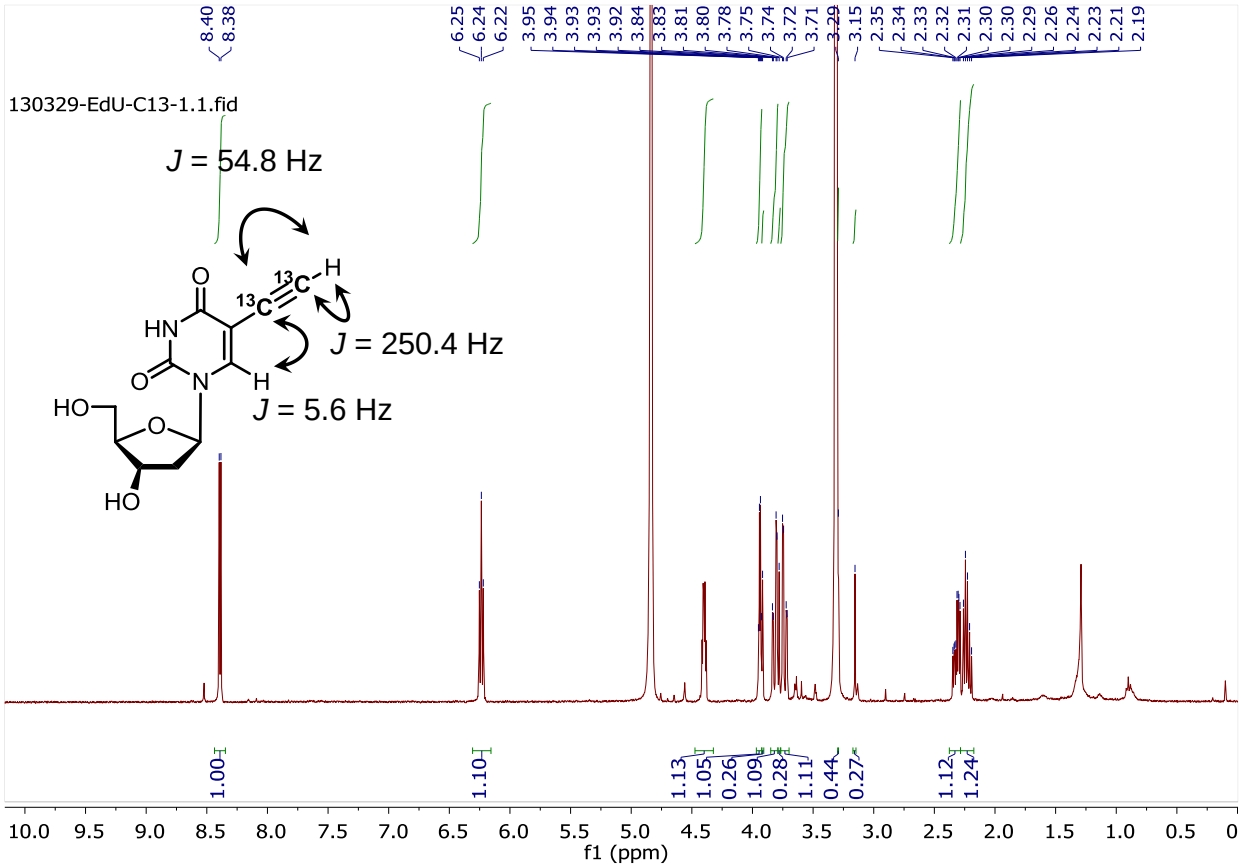
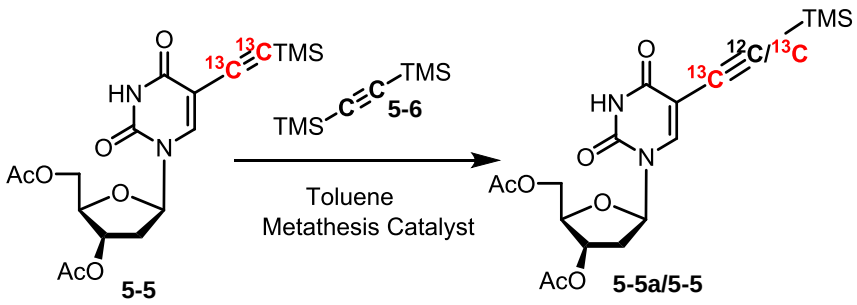
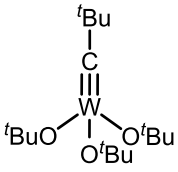



**Figure 5-S1. $^{1}$H NMR (400 MHz) of 5-3 in $CD_3OD$.**

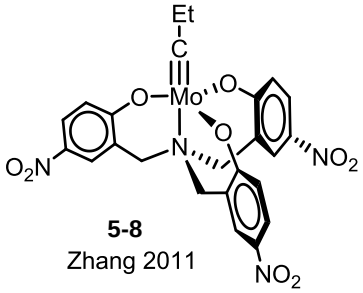
Investigation of conditions for the alkyne cross-metathesis (Table 5-1).




Catalysts:



**5-7**
Schrock 1981



**5-8**
Zhang 2011

| Entry | Catalyst | **5-6** loading | Temp. | Time | Yield | **5-5a:5-5** |
|---|---|---|---|---|---|---|
| 1 | **5-7**, 5 eq. | 100 eq. | RT | 12h | N/R | |
| 2 | **5-7**', 5 eq. | 100 eq. | 80°C | 12h | Side Product | |
| 3 | **5-8**, 5 eq. | 100 eq. | RT | 12h | N/R | |
| 4 | **5-8**, 5 eq. | 100 eq. | 80°C | 12h | 20% | 2.2 : 1 |
| 5 | **5-8**, 0.5 eq. | 10 eq. | 80°C | 12h | N/R | |

General procedure:

In a glove box filled with Ar, substrate **5-6** and catalyst **5-7/5-8** was added to a solution of **5-5** in 0.5 mL toluene (catalyst **5-8** was prepared in $CCl_4$ *in situ* according to Jyothish and Zhang (*Angew. Chem. Int. Ed. Engl.* **50,** 3435–8 (2011))). The mixture was left at the indicated temperature for 12 h before concentrated *in vacuo*. The residue was purified by column chromatography on silica gel (0-70% Ethyl acetate in Hexanes) to recover **5-5** or to yield **5-5/5-5a**.

Entry 2: The starting material is not recovered. A side product, tentatively assigned as the C1' anomer of **5-5** according to $^1$H NMR (Figure 5-S2), is isolated.

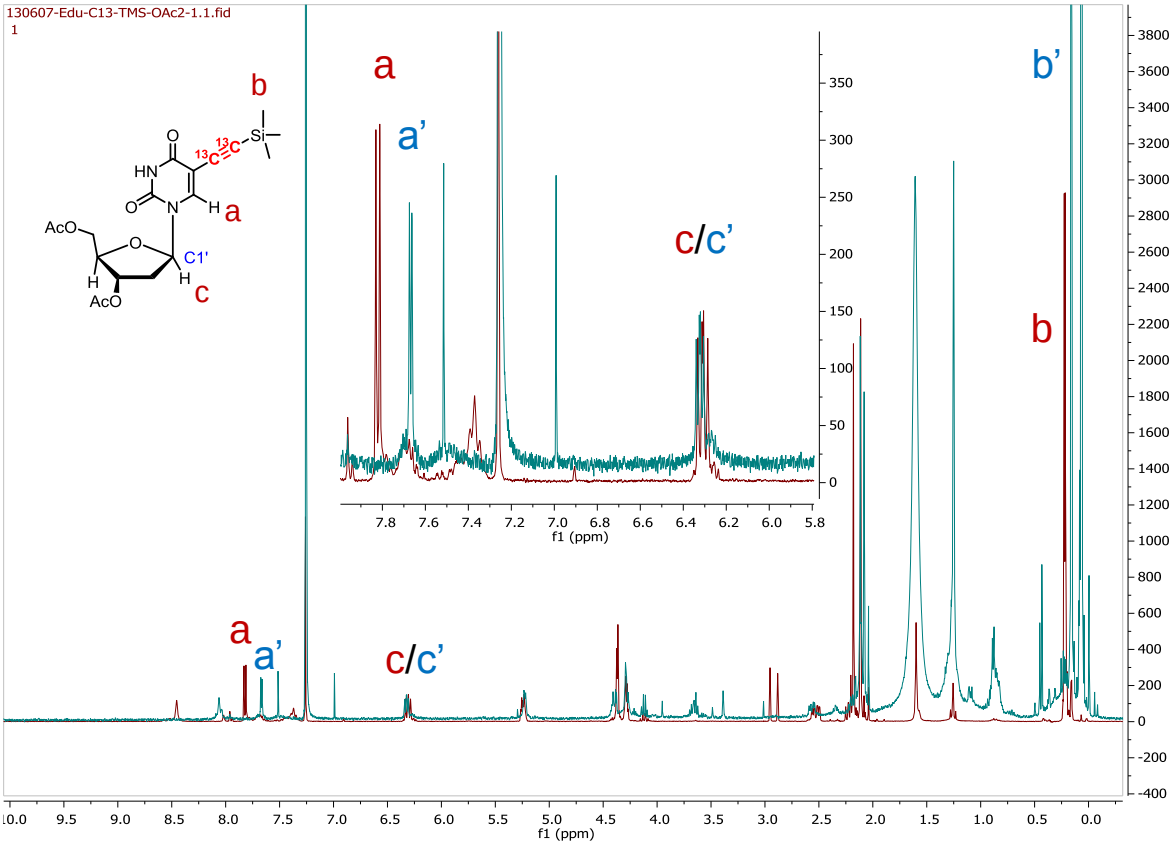



**Figure 5-S2. $^{1}$H NMR spectra of 5-5 (red) and the side product isolated (cyan).**

Entry 4: A mixture of **5-5** and **5-5a** is isolated (yield 20%). The mixture is characterized by $^{1}$H NMR (Figure 5-S3). The isotopic ratio is determined according to peak intensities of HRMS (Figure 5-S4).

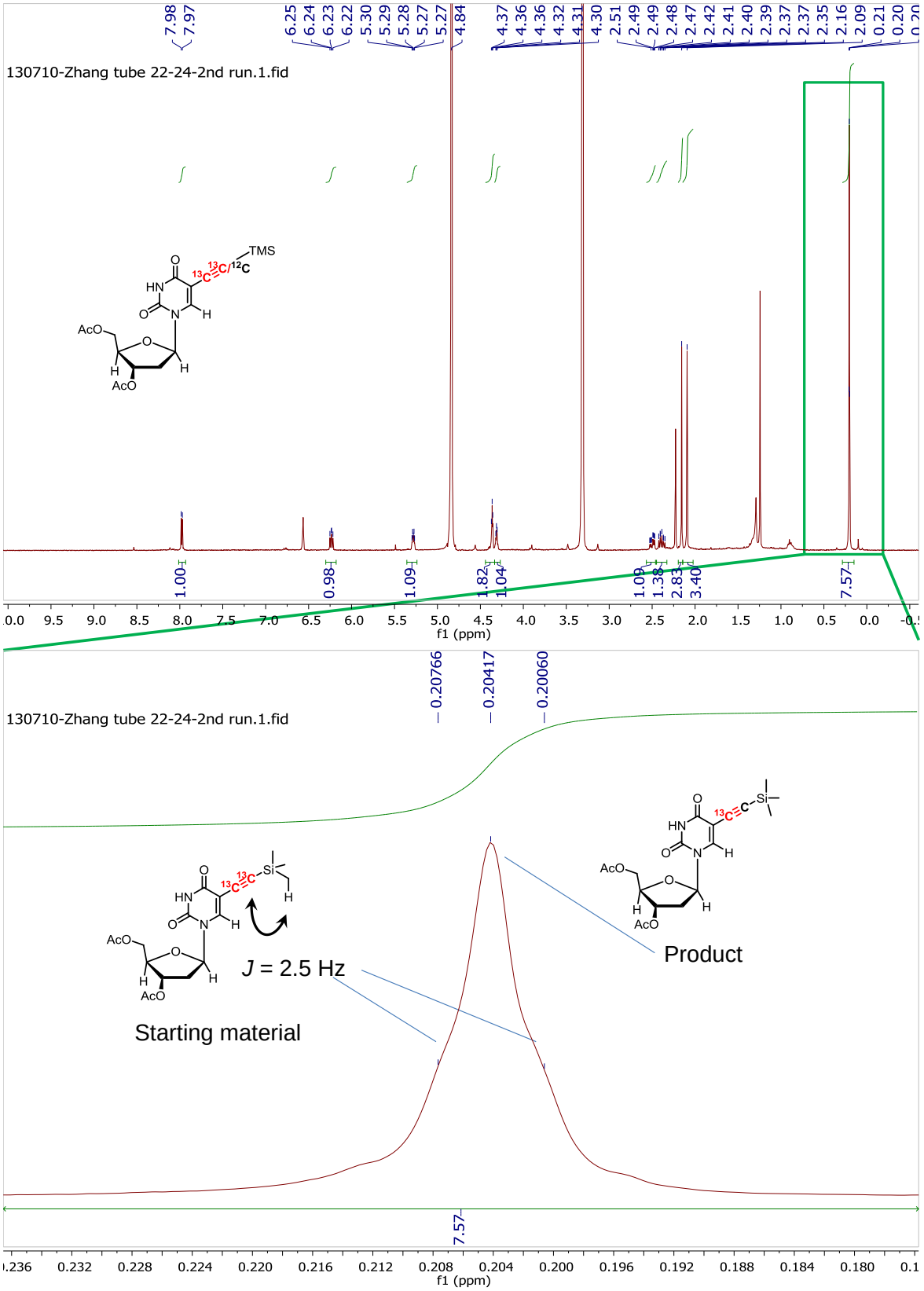



**Figure 5-S3. $^{1}$H NMR (400 MHz) of mixture 5-5/5-5a in $CD_3OD$.**

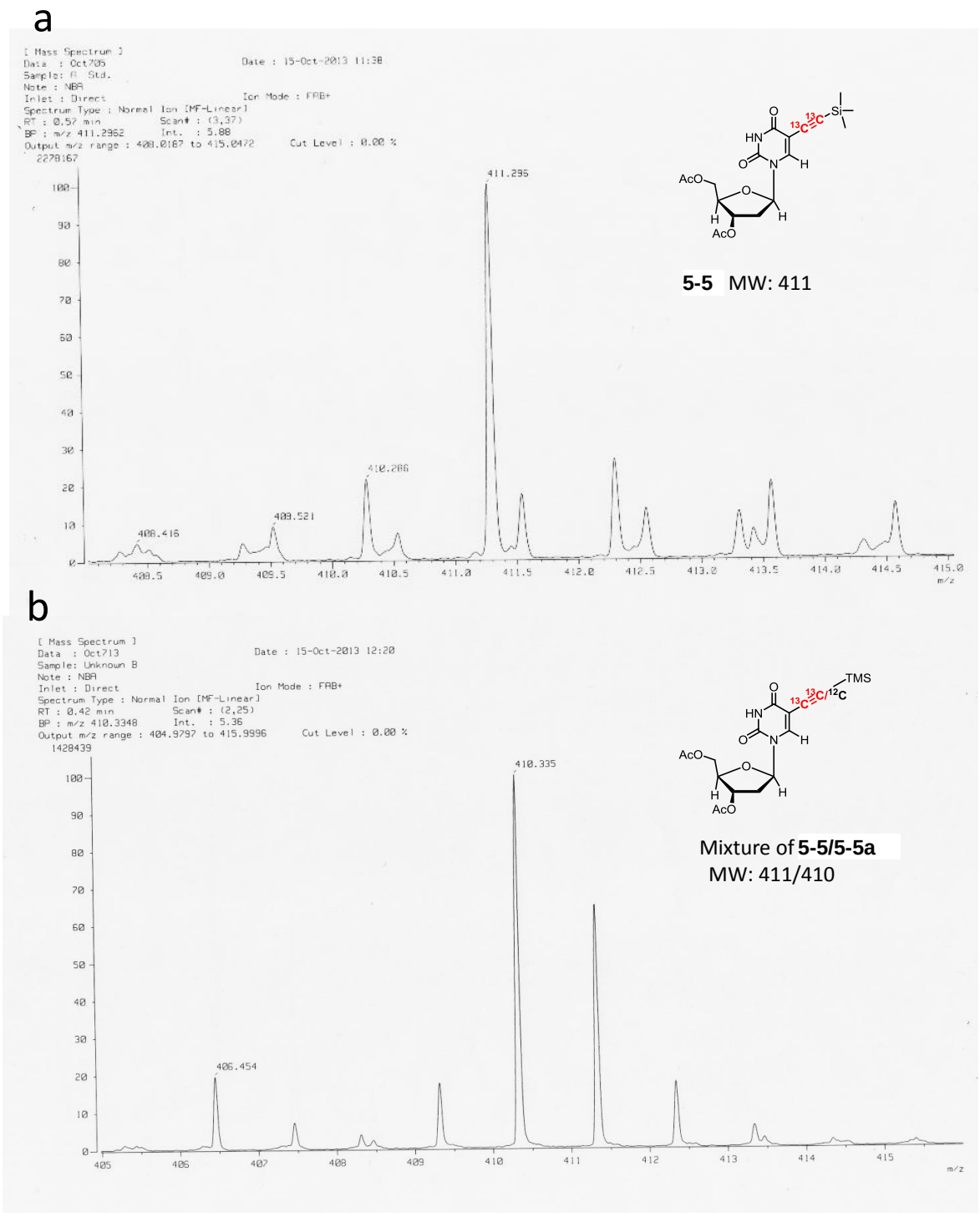
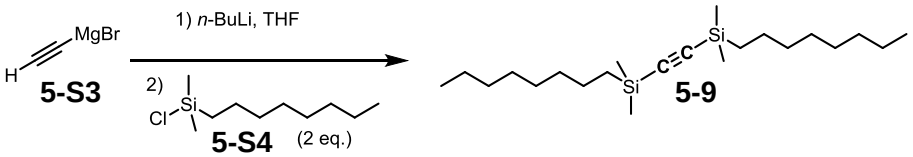



**Figure 5-S4. HRMS (FAB+) of (a) 5 and (b) mixture of 5-5/5-5a.**

Synthesis of **5-9**:

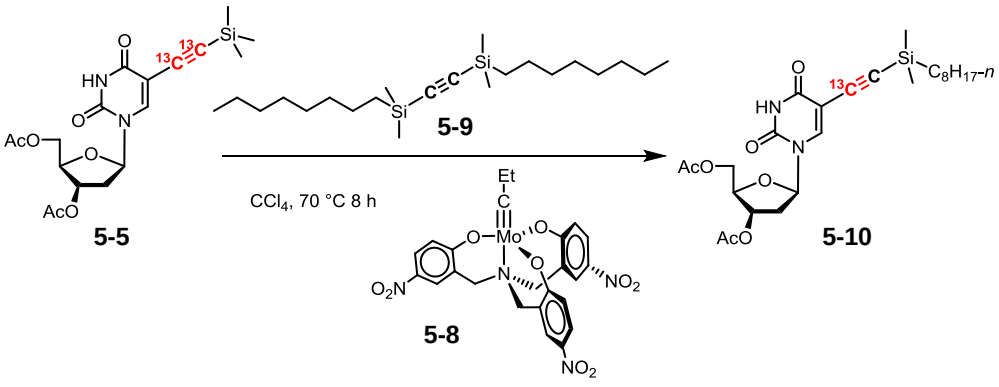



To a solution of ethynylmagnesium bromide in THF (5.0 ml, 0.5 M solution, 2.5 mmol) was added 15 ml THF under Ar. The solution was cooled to -78 ℃ and 2.4 ml *n*-BuLi in hexane (1.6 M, 3.8 mmol) was added dropwisely. After 30 min, chloro(dimethyl)octylsilane (1.21 ml, 1.06 g, 5.1 mmol) was added dropwisely. The reaction was then warmed to RT and stirred for another 3 h before filtered through a short pad of silica. The solvent was removed under reduced pressure and the residue was purified by column chromatography on silica gel (pure Hexanes) to give **5-9** (885 mg, 2.4 mmol, 96%) as a colorless liquid.

$^{1}$H NMR (400 MHz, Chloroform-*d*) δ 1.42 – 1.24 (m, 24 H), 0.88 (t, *J* = 6.6 Hz, 6 H), 0.60 (dd, *J* = 9.4, 6.2 Hz, 4 H), 0.13 (s, 12 H). $^{13}$C NMR (101 MHz, Chloroform-*d*) δ 113.94, 33.37, 32.12, 29.49, 29.43, 23.92, 22.85, 16.26, 14.27, -1.55.

HRMS (EI+) m/z Calcd. for $C_{22}H_{46}Si_2$ $[M]^+$:366.3138. Found: 366.3134

Synthesis of **5-10**:

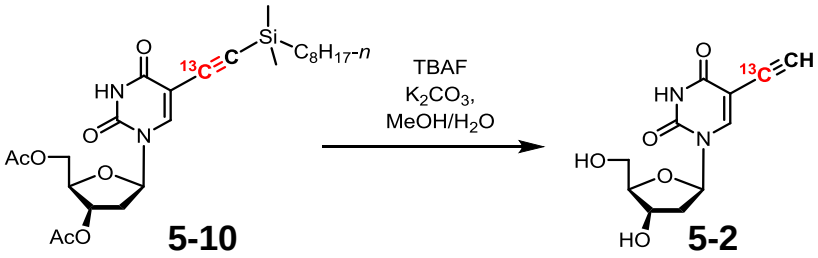


In a glove box filled with Ar, catalyst **5-8** (36.5 μmol, 5 eq.) was prepared in 0.5 mL dry $CCl_4$ *in situ* according to the procedure documented by Jyothish and Zhang (*Angew. Chem. Int. Ed. Engl.* **50,** 3435–8 (2011)). To the solution of **5-8** in $CCl_4$ was added 9 (267 mg, 0.73 mmol) and a solution of **5-5** (3.0 mg, 7.3 μmol) in 0.5 mL dry $CCl_4$. The mixture was heated to 70 ℃ for 8 h before concentrated *in vacuo*. The residue was purified by column chromatography on silica gel (0-70% Ethyl acetate in Hexanes) to recover **5-5** (0.5 mg, 1.2 μmol) and to give **5-10** (1.0 mg, 2.0 μmol, 27%, 33% B.R.S.M.) as a thin film.

$^1$H NMR (400 MHz, Methanol-$d_4$) δ 7.97 (d, $J$ = 5.6 Hz, 1 H), 6.23 (dd, $J$ = 7.9, 5.9 Hz, 1 H), 5.28 (dt, $J$ = 6.8, 2.4 Hz, 1 H), 4.36 (dd, $J$ = 5.8, 3.4 Hz, 2 H), 4.31 (dd, $J$ = 6.3, 3.2 Hz, 1 H), 2.50 (ddd, $J$ = 14.5, 6.1, 2.5 Hz, 1 H), 2.38 (ddd, $J$ = 20.2, 7.7, 6.1 Hz, 1 H), 2.15 (s, 3 H), 2.09 (s, 3 H), 1.30 (s, 12 H), 0.90 (t, $J$ = 6.9 Hz, 3 H), 0.70 – 0.62 (m, 2 H), 0.18 (s, 6 H). $^{13}$C NMR (101 MHz, $CD_3OD$) δ 97.56.

MS (FAB+) m/z Calcd. for $C_{24}{}^{13}CH_{38}N_2NaO_7Si$ $[M+Na]^+$: 530.24. Found: 530.25

Synthesis of **5-2**:




To a solution of **5-10** (0.4 mg, 0.8 μmol) in 0.5 ml MeOH and 0.05 ml $H_2O$ was added $K_2CO_3$ (2.0 mg, 14 μmol) and TBAF (20 μL, 1 M in THF) at RT. The reaction was stirred 7 h at RT before concentrated *in vacuo*. The residue was purified by reverse phase HPLC to give compound **5-2** (0.1 mg, 0.4 μmol, ~50%) as a thin film.

HPLC condition: 20 min gradient elution using $H_2O$:MeCN starting from 100:0 to 85:15. Retention time: 15.4 min

The mass of the product is determined by UV-Vis ($\lambda_{abs}$ = 288 nm, ε = 12,000 $cm^{-1}M^{-1}$ in methanol).

$^1$H NMR (500 MHz, Methanol-$d_4$) δ 8.39 (d, $J$ = 5.7 Hz, 1 H), 6.24 (t, $J$ = 6.5 Hz, 1 H), 4.40 (dt, $J$ = 6.6, 3.6 Hz, 1 H), 3.94 (q, $J$ = 3.3 Hz, 1 H), 3.82 (dd, $J$ = 12.0, 3.1 Hz, 1 H), 3.73 (dd, $J$ = 12.0, 3.4 Hz, 1 H), 3.53 (d, $J$ = 51.3 Hz, 1 H), 2.32 (ddd, $J$ = 13.6, 6.2, 3.7 Hz, 1 H), 2.27 – 2.17 (m, 1 H). $^{13}$C NMR (101 MHz, $CD_3OD$) δ 76.00. MS (ESI+) m/z Calcd. for $C_{10}{}^{13}CH_{13}N_2O_5$ $[M+H]^+$: 254.09. Found: 254.70

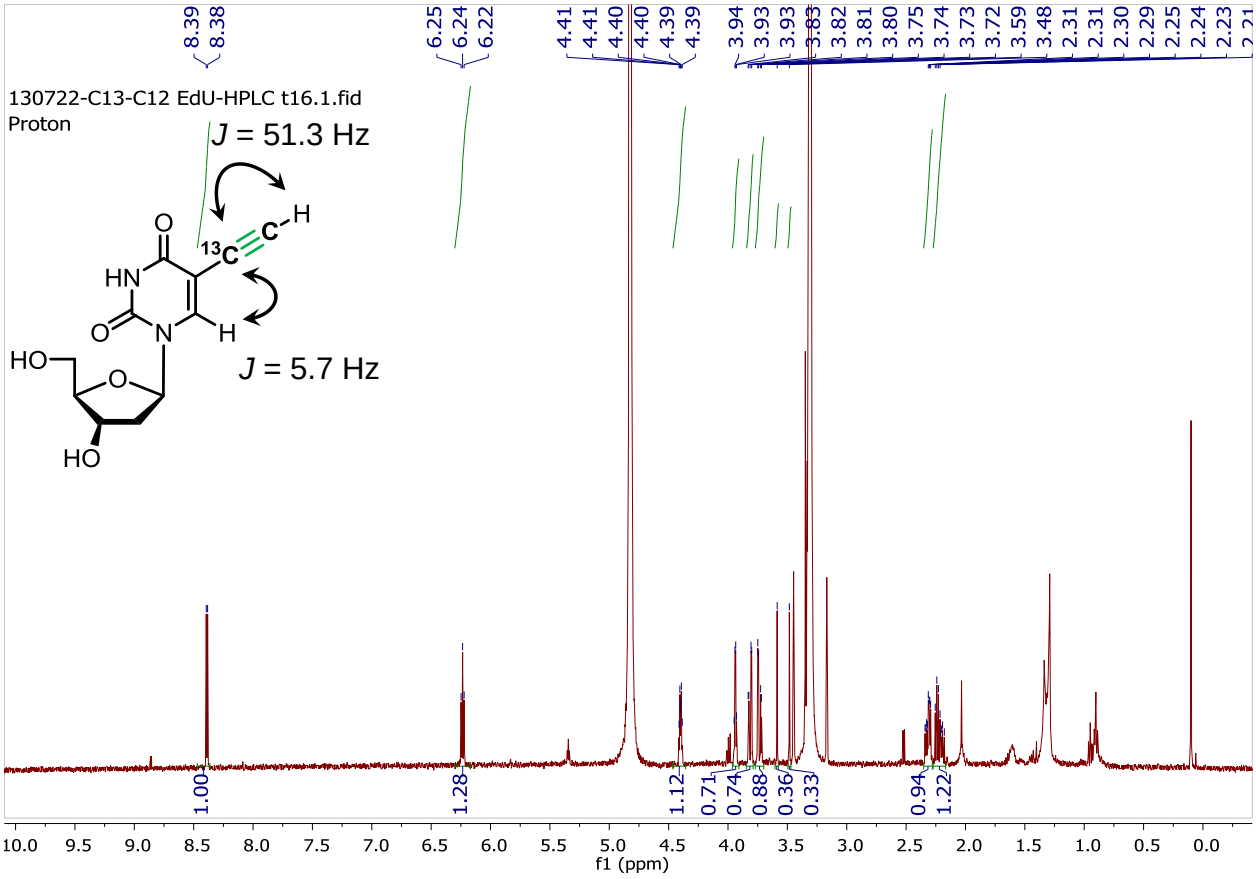
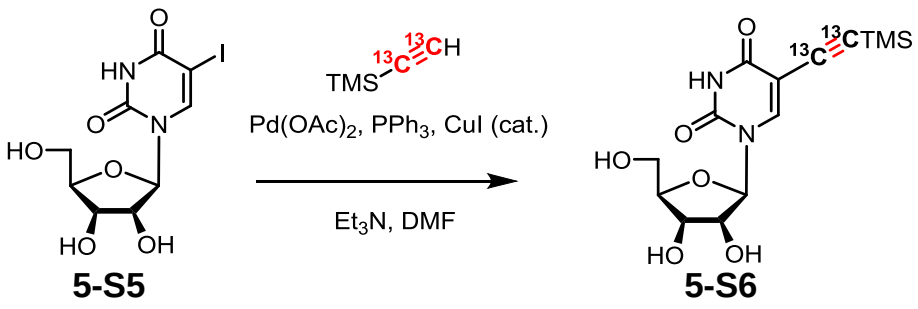



**Figure 5-S5. $^1$H NMR (500 MHz) of 5-2 in $CD_3OD$.**

Synthesis of **5-S6**:

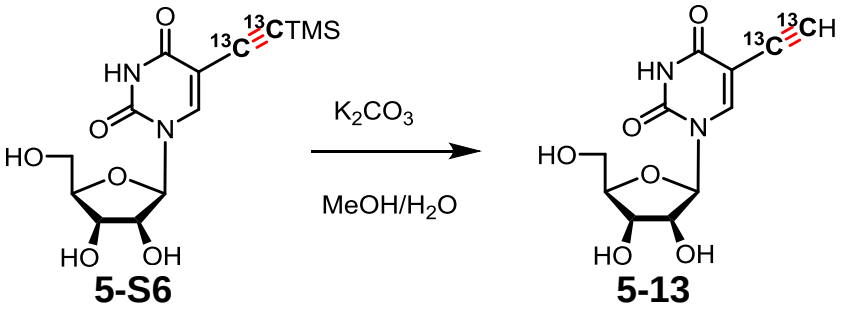



To an oven-dried vial was added **5-S5** (15 mg, 50 μmol), $Pd(OAc)_2$ (1.1 mg, 5 μmol), $PPh_3$ (2.6 mg, 10 μmol) CuI (1.0 mg, 5 μmol), DMF (1 ml), $Et_3N$ (15 mg, 20.7 μl, 150 μmol) and TMS$^{13}C\equiv{}^{13}CH$ (7.5 mg, 10.8 μl, 75 μmol) under Ar. The mixture was stirred at RT for 12 h before concentrated *in vacuo*. The residue was purified by column chromatography on silica gel (0-50% methanol in dichloromethane) to give **5-S6** (9.0 mg, 26 μmol, 52%) as a thin film.

$^1$H NMR (400 MHz, Methanol-$d_4$) δ 8.41 (d, $J$ = 4.9 Hz, 1 H), 5.91 – 5.83 (m, 1 H), 4.21 – 4.13 (m, 2 H), 4.07 – 3.98 (m, 1 H), 3.88 (dd, $J$ = 12.2, 2.6 Hz, 1 H), 3.75 (dd, $J$ = 12.2, 2.8 Hz, 1 H), 0.20 (d, $J$ = 2.3 Hz, 9 H). $^{13}$C NMR (101 MHz, $CD_3OD$) δ 99.24 (d, $J$ = 141.0 Hz), 96.95 (d, $J$ = 141.0 Hz).

MS (FAB+) m/z Calcd. for $C_{12}{}^{13}C_2H_{21}N_2O_6Si$ $[M+H]^+$: 343.12. Found: 343.17

Synthesis of **5-13**:




To a solution of **5-S6** (3.0 mg, 8.8 μmol) in 0.6 ml MeOH and 0.1 ml $H_2O$ was added $K_2CO_3$ (5.0 mg, 36 μmol) at RT. The reaction was stirred overnight before concentrated *in vacuo*. The residue was purified by reverse phase HPLC to give compound **5-13** (2.2 mg, 8.1 μmol, 92%) as a thin film.

$^1$H NMR (400 MHz, Methanol-$d_4$) δ 8.47 (d, $J$ = 5.6 Hz, 1 H), 5.93 – 5.83 (m, 1 H), 4.21 – 4.13 (m, 2 H), 4.06 – 3.98 (m, 1 H), 3.88 (dd, $J$ = 12.2, 2.6 Hz, 1 H), 3.75 (dd, $J$ = 12.2, 2.8 Hz, 1H), 3.54 (dd, $J$ = 250.4, 54.6 Hz). $^{13}$C NMR (101 MHz, $CD_3OD$) δ 82.90 (d, $J$ = 180.2 Hz), 75.74 (d, $J$ = 180.2 Hz).

MS (ESI+) m/z Calcd. for $C_9{}^{13}C_2H_{13}N_2O_6$ $[M+H]^+$: 271.08. Found: 271.51

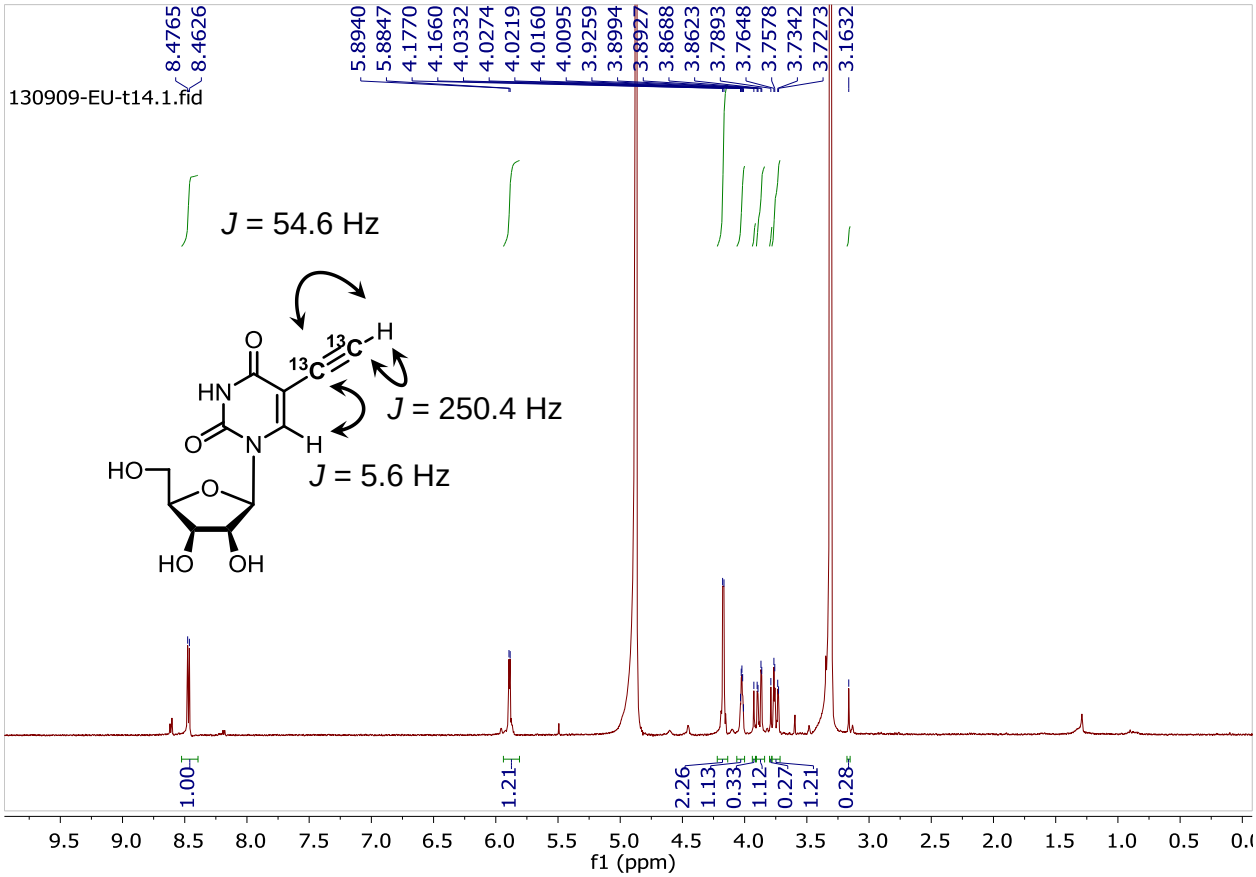



**Figure 5-S6. $^{1}$H NMR (500 MHz) of 5-2 in $CD_3OD$.**

## 5.8 References


1. Lavis, L. D.; Raines, R. T. *ACS Chem. Biol.* **2008**, *3*, 142.
2. Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. *Nat. Methods* **2005**, *2*, 905.
3. Michalet, X.; Pinaud, F. F.; Bentolila, L. A.; Tsay, J. M.; Doose, S.; Li, J. J.; Sundaresan, G.; Wu, A. M.; Gambhir, S. S.; Weiss, S. *Science* **2005**, *307*, 538.
4. Miyawaki, A. *Annu. Rev. Biochem.* **2011**, *80*, 357.
5. Donnert, G.; Keller, J.; Wurm, C. A.; Rizzoli, S. O.; Westphal, V.; Schönle, A.; Jahn, R.; Jakobs, S.; Eggeling, C.; Hell, S. W. *Biophys. J.* **2007**, *92*, L67.
6. Bates, M.; Huang, B.; Dempsey, G. T.; Zhuang, X. *Science* **2007**, *317*, 1749.
7. Zhao, Y.; Araki, S.; Wu, J.; Teramoto, T.; Chang, Y.-F.; Nakano, M.; Abdelfattah, A. S.; Fujiwara, M.; Ishihara, T.; Nagai, T.; Campbell, R. E. *Science* **2011**, *333*, 1888.

8. Yan, P.; Acker, C. D.; Zhou, W.-L.; Lee, P.; Bollensdorff, C.; Negrean, A.; Lotti, J.; Sacconi, L.; Antic, S. D.; Kohl, P.; Mansvelder, H. D.; Pavone, F. S.; Loew, L. M. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20443.
9. Wei, L.; Yu, Y.; Shen, Y.; Wang, M. C.; Min, W. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 11226.
10. Yamakoshi, H.; Dodo, K.; Okada, M.; Ando, J.; Palonpon, A.; Fujita, K.; Kawata, S.; Sodeoka, M. *J. Am. Chem. Soc.* **2011**, *133*, 6102.
11. Yamakoshi, H.; Dodo, K.; Palonpon, A.; Ando, J.; Fujita, K.; Kawata, S.; Sodeoka, M. *J. Am. Chem. Soc.* **2012**, *134*, 20681.
12. Wei, L.; Hu, F.; Shen, Y.; Chen, Z.; Yu, Y.; Lin, C.-C.; Wang, M. C.; Min, W. *Nat. Methods* **2014**, *11*, 410.
13. Freudiger, C. W.; Min, W.; Saar, B. G.; Lu, S.; Holtom, G. R.; He, C.; Tsai, J. C.; Kang, J. X.; Xie, X. S. *Science* **2008**, *322*, 1857.
14. Saar, B. G.; Freudiger, C. W.; Reichman, J.; Stanley, C. M.; Holtom, G. R.; Xie, X. S. *Science* **2010**, *330*, 1368.
15. Min, W.; Freudiger, C. W.; Lu, S.; Xie, X. S. *Annu. Rev. Phys. Chem.* **2011**, *62*, 507.
16. Cheng, J.-X.; Xie, X. S. *Coherent Raman Scattering Microscopy*; CRC Press, 2012.
17. Liu, Z.; Li, X.; Tabakman, S. M.; Jiang, K.; Fan, S.; Dai, H. *J. Am. Chem. Soc.* **2008**, *130*, 13540.
18. Paul, C.; Wang, J.; Wimley, W. C.; Hochstrasser, R. M.; Axelsen, P. H. *J. Am. Chem. Soc.* **2004**, *126*, 5843.
19. Dieringer, J. A.; Lettan, R. B.; Scheidt, K. A.; Van Duyne, R. P. *J. Am. Chem. Soc.* **2007**, *129*, 16249.
20. Etchegoin, P. G.; Le Ru, E. C.; Meyer, M. *J. Am. Chem. Soc.* **2009**, *131*, 2713.
21. Silverman, L. N.; Pitzer, M. E.; Ankomah, P. O.; Boxer, S. G.; Fenlon, E. E. *J. Phys. Chem. B* **2007**, *111*, 11611.
22. Lipkin, J. S.; Song, R.; Fenlon, E. E.; Brewer, S. H. *J. Phys. Chem. Lett.* **2011**, *2011*, 1672.
23. Fürstner, A. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 2794.
24. Baker, J.; Jarzecki, A. A.; Pulay, P. *J. Phys. Chem. A* **1998**, *102*, 1412.
25. Salic, A.; Mitchison, T. J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 2415.
26. Yu, C.-S.; Oberdorfer, F. *Synlett* **2000**, *2000*, 86.
27. Katz, T. J.; McGinnis, J. *J. Am. Chem. Soc.* **1975**, *97*, 1592.
28. Schrock, R. R. *Acc. Chem. Res.* **1986**, *19*, 342.
29. Haberlag, B.; Freytag, M.; Daniliuc, C. G.; Jones, P. G.; Tamm, M. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 13019.

30. McCullough, L. G.; Listemann, M. L.; Schrock, R. R.; Churchill, M. R.; Ziller, J. W. *J. Am. Chem. Soc.* **1983**, *105*, 6729.

31. Fürstner, A.; Mathes, C. *Org. Lett.* **2001**, *3*, 221.

32. Wengrovius, J. H.; Sancho, J.; Schrock, R. R. *J. Am. Chem. Soc.* **1981**, *103*, 3932.

33. Hickmann, V.; Alcarazo, M.; Fürstner, A. *J. Am. Chem. Soc.* **2010**, *132*, 11042.

34. Jyothish, K.; Zhang, W. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 3435.

35. Alfano, R. R.; Baldeck, P. L.; Ho, P. P.; Agrawal, G. P. *J. Opt. Soc. Am. B* **1989**, *6*, 824.

36. Grammel, M.; Hang, H. C. *Nat. Chem. Biol.* **2013**, *9*, 475.

37. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 2004.

38. Jao, C. Y.; Salic, A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 15779.

39. Martin, B. R.; Cravatt, B. F. *Nat. Methods* **2009**, *6*, 135.

40. Heppekausen, J.; Stade, R.; Goddard, R.; Fürstner, A. *J. Am. Chem. Soc.* **2010**, *132*, 11045.

41. Yang, H.; Liu, Z.; Zhang, W. *Adv. Synth. Catal.* **2013**, *355*, 885.

42. Paley, D. W.; Sedbrook, D. F.; Decatur, J.; Fischer, F. R.; Steigerwald, M. L.; Nuckolls, C. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 4591.

43. Lin, Y. A.; Boutureira, O.; Lercher, L.; Bhushan, B.; Paton, R. S.; Davis, B. G. *J. Am. Chem. Soc.* **2013**, *135*, 12156.

44. Kurhanewicz, J.; Bok, R.; Nelson, S. J.; Vigneron, D. B. *J. Nucl. Med.* **2008**, *49*, 341.

45. Senyo, S. E.; Steinhauser, M. L.; Pizzimenti, C. L.; Yang, V. K.; Cai, L.; Wang, M.; Wu, T.-D.; Guerquin-Kern, J.-L.; Lechene, C. P.; Lee, R. T. *Nature* **2013**, *493*, 433.

46. Palonpon, A. F.; Ando, J.; Yamakoshi, H.; Dodo, K.; Sodeoka, M.; Kawata, S.; Fujita, K. *Nat. Protoc.* **2013**, *8*, 677.

47. Palonpon, A. F.; Sodeoka, M.; Fujita, K. *Curr. Opin. Chem. Biol.* **2013**, *17*, 708.

48. Xu, C.; Wise, F. W. *Nat. Photonics* **2013**, *7*, 875.

49. Zhang, D.; Slipchenko, M. N.; Leaird, D. E.; Weiner, A. M.; Cheng, J.-X. *Opt. Express* **2013**, *21*, 13864.

50. Kong, L.; Ji, M.; Holtom, G. R.; Fu, D.; Freudiger, C. W.; Xie, X. S. *Opt. Lett.* **2013**, *38*, 145.

51. Sletten, E. M.; Bertozzi, C. R. *Acc. Chem. Res.* **2011**, *44*, 666.

52. Debets, M. F.; van Berkel, S. S.; Dommerholt, J.; Dirks, A. T. J.; Rutjes, F. P. J. T.; van Delft, F. L. *Acc. Chem. Res.* **2011**, *44*, 805.

53. Devaraj, N. K.; Weissleder, R. *Acc. Chem. Res.* **2011**, *44*, 816.

54. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046.

55. Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518.